

Open Research Online

The Open University's repository of research publications
and other research outputs

Moral Luck in Medical Ethics and Practical Politics

Thesis

How to cite:

Dickenson, Donna (1989). Moral Luck in Medical Ethics and Practical Politics. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1989 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000fc33>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

DX88774

UNRESTRICTED

Moral Luck in Medical Ethics and Practical Politics

Donna Dickenson, B.A., M.Sc. (Econ.)

Submitted for the Degree of Doctor of Philosophy

Department of Philosophy

Faculty of Arts

June 1989

No part of this material has previously been submitted for a degree or other qualification at the Open University or any other institution.

Date of submission : 29th June 1989

Date of award : 4th October 1989

ProQuest Number: 27758699

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 27758699

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All Rights Reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Moral Luck in Medical Ethics and Practical Politics

Abstract

Typically we maintain two incompatible standards towards right action and good character, and the tension between these polarities creates the paradox of moral luck. In practice we regard actions as right or wrong, and character as good or bad, partly according to what happens as a result of the agent's decision. Yet we also think that people should not be held responsible for matters beyond their control.

This split underpins Kant's assertion that only the good will is securely good, that its goodness is impervious to outcome ill-luck. Some commentators, such as Martha Nussbaum and to some extent Bernard Williams, think that this simply writes off the paradox. Williams asserts that the paradox is insoluble, and that its inescapability threatens the notion of agent responsibility. In contrast Thomas Nagel argues that agents' most cherished projects may be indeed be subject to luck, but that does not mean that their deepest motivations are moral. This, I suggest, is one of several means whereby we might limit the effect of the paradox without denying that the tension exists. But I also argue that it is wrong to accuse Kant of ignoring the paradox.

Ethical consequentialists, on the other hand, appear to have no problem with moral luck, because the paradox depends on a dichotomy between the outside world and the locus of moral worth in the individual agent. But this turns out not to be true. The problem of moral luck is not some strange Kantian fixation, but a general dilemma: a variant on what

Nagel terms "the problem of excess objectivity" which cuts across all of ethics and metaphysics.

Retaining a broadly Kantian notion of agent-responsibility, but limiting what agents are responsible for, requires us to delineate the realm of ethics more narrowly than has been done by those who believe that the rational and/or prudential are coterminous with the ethical. This strategy for minimising the paradox's impact is explored in two areas from medical ethics, the allocation of scarce medical resources and informed consent, and two from public policy, secrecy and nuclear deterrence.

Throughout, the analysis seeks to test Nagel's maxim that the best we can hope for is to act in such a manner that we would not have to revise our opinion of how we should have acted once the consequences of our actions become apparent.

TABLE OF CONTENTS

| | |
|---|---------|
| 1. Ethics versus Luck? | 1 |
| 2. Risk and Kant | 27 |
| 3. Probability and Utilitarianism | 58 |
| 4. Rationality and Prudence | 85 |
| 5. Moral Luck and the Allocation of Scarce Medical Goods | 116 |
| 6. Secrecy and Moral Luck | 145 |
| 7. Responsibility, Risk, Rationality and Consent | 175 |
| 8. The Illest Chance? Luck, Moral Mathematics and Nuclear Deterrence | 204 |
| Bibliography | 242 |

Chapter One

Ethics versus Luck?

Is "moral" incompatible with "luck"? The attempt to shore up the ethically right against the tidal waves of ill chance is generally thought central to the Kantian moral enterprise. But if that effort is hopelessly quixotic, or even misguided, is Kantianism in danger? More broadly, is ethics as a whole somehow at risk?

Typically, we maintain two incompatible standards towards right actions and good character, and the tension between these polarities creates the paradox of moral luck. In practice we regard actions as right or wrong, and moral character as good or bad, partly according to what happens as a result of the agent's decision. That is, we make responsibility hinge to some extent on things outside the agent's control. Yet at the same time we think that people should not be held responsible for matters beyond their control.

This tension underpins Kant's famous assertion that only the good will is securely good, and that its goodness is impervious to ill-luck in how things actually turn out.

Even if it should happen that by a particularly unfortunate fate or by the niggardly provision of a stepmotherly nature, this will should be wholly lacking in power to accomplish its purpose, if even the greatest effort should not avail it to achieve anything of its end, and if there remained only the good will (not as a mere wish, but as the summoning of all the means in our power), it would sparkle like a

jewel in its own right, as something that had its full worth in itself.¹

But is moral luck only a problem for a Kantian? Martha Craven Nussbaum has claimed that not only Kant but also Plato was motivated by a concern to minimise the effects of chance on moral character and the rightness of ethical choices.² And though ethical consequentialism appears not to have a problem with moral luck, in fact it does so, as I shall argue in chapter three. But most particularly, I want to see whether the concept of moral luck can cast any light on problems in medical ethics and practical politics--and conversely, how grounding the concept in practical ethics might help us to better understand and perhaps resolve the paradox.

This moral luck debate, begun by Bernard Williams and Thomas Nagel in 1976,³ has become reasonably familiar to many ethicists. And its application has been extended to theories outside the original scope of the debate, most notably in Nussbaum's work--which stretches the concept to fit classical tragedy as well as philosophy. But it has not yet been applied to practical ethics, and this is what I intend to do in chapters five through eight. The areas which I discuss--the allocation of scarce medical resources, informed consent, secrecy in politics, and nuclear deterrence--can be fruitfully explored, I argue, in terms of the simple question arising from a concern with moral luck: what happens if things go wrong? This practical ethics section occupies slightly more than half of the dissertation. In it I conclude that a modified Kantian approach can handle the moral luck problem as it surfaces in these areas of medical ethics and practical politics.

In the first four chapters I erect the theoretical framework for what I hope will be a more sophisticated discussion than might be suggested by the basic question of what happens when things beyond an agent's control go wrong. I begin in this chapter with a systematic exposition of the moral luck debate between Williams and Nagel, considering also the ramifications of the debate in further articles on the same subject by other writers⁴. (It has also surfaced in such unexpected areas as feminism.⁵) I end with Williams's own further reflections on the significant debate which he set in train.⁶

The second chapter analyses the Kantian sources of the claim that ethics cannot allow itself to be undermined by chance, and deals at some length with Nussbaum's stimulating critique of what she presents as Kant's view as well as Plato's.⁷ It asks whether Kant's approach to the moral luck paradox is essentially sound, despite Nussbaum's criticisms, and concludes in part that Nussbaum has misrepresented Kant's solution, which is closer to her own than she acknowledges. In chapter three I turn from deontology to consequentialism, and particularly to utilitarianism. Here I explore the threat from moral luck, which I argue is a threat even to consequentialists, in terms of actual and potential consequences, with the associated concept of probability also entering in. In the course of chapter three I suggest that the best way to resolve the paradox turns out to be retaining a broadly Kantian conception of agency, with considerable strictness about responsibility, but limiting severely what agents are responsible for. This requires an equally strict delineation of where the ethical ends and the merely prudential begins, and this I attempt to do in

chapter four, which also considers rationality. Thus each of chapters two through four also introduces a new but associated topic arising from moral luck: risk in chapter two, probability in chapter three, and rationality in chapter four.

In this chapter I summarise the Williams-Nagel debate and ask whether the moral luck paradox is real. But first I need to provide the first of many qualifying statements about luck and ethics. Not all "luck" is "moral", but more of the ethical is open to the operations of chance than we think, or perhaps want to think. If I give ten pounds to Oxfam as an expression of my belief that it is ethically correct to feed the hungry, I do not regard my decision as having been proved unethical by a flash flood which stops the supplies that I have donated from getting through.⁸ But I am, of course, lucky to have ten pounds to give, and my generosity is somehow slightly less than my own: it is partly a matter of my good fortune in living in a comparatively wealthy country and in being in work. (Of course, many others also live in a First World country and hold paid employment, but do not contribute to charity--so my generosity is partly my own.) In a further sense, I am lucky to live under a very watered-down version of Christianity which no longer enjoins me to sell all I have and give my goods to the poor. I can be accounted generous only because the moral standards set by my society are actually rather lax in the matter of financial altruism, and in this sense I am again lucky to live in this society.

These last two points are applications of what Nagel calls luck in one's antecedent circumstances and luck in the problems which have to be faced.⁹ More commonly, moral luck is considered in relation to luck in outcomes, and this is the sense in which I (and Williams) use it most of the time. A recurring theme in this dissertation is that the extent of outcome luck has been underestimated. Ethicists have often failed to consider all the possible outcomes or have even assumed that the desired outcome is automatic. Thus, to take the most hackneyed example, the old saw about whether it is ethically justifiable to kill one person in order to save one hundred almost always presumes that the hundred will be saved by the death of the one. Even Williams himself can be accused of underestimating the number of possible outcomes in his well-known hypothetical case of Pedro and Jim.¹⁰

In this example a traveller, Jim, happens into a South American market square where a captain, Pedro, is about to execute twenty Indians. Pedro offers Jim a chance to save all the Indians but one, whom Jim must kill with his own hand. If Jim refuses, Pedro says that all the executions will proceed; if he assents, Pedro promises that the other nineteen Indians will go free. Williams concludes rather grudgingly that the utilitarian prescription--that Jim should kill the Indian--is right in this case, though not for conventional utilitarian reasons: "...[T]he utilitarian is probably right in this case [but] that is not to be found out just by asking the utilitarian's questions."¹¹ Why Williams comes down in favour of Jim's killing the Indian is not a concern of my argument; I merely wish to point out that he has not fully considered what happens if things go wrong.

It is all too easy--the stuff of B-grade movies--to imagine Jim turning trembling to Pedro, smoking revolver in hand, and the camera panning onto Pedro's delighted grin as he takes the gun and shoots the first of the nineteen other Indians whom he meant to kill all along.

Does this "unlucky" outcome make Jim's decision wrong?¹² A Kantian would assert that it was wrong all along, regardless of the outcome. More importantly, a Kantian could assert that its rightness or wrongness is independent of the outcome. Now possibly some consequentialists would argue that no additional diminution of welfare results from Jim's decision, since the worst would happen anyway: all twenty Indians would have died in any case. But having failed to consider all the consequences should be a cause of self-mortification for any utilitarian strategy which incorporates probability assessments. (This would include expected-value strategy, the most common decision-theoretic model, which advises the agent to choose among courses of action by multiplying the probability that an action will produce a desired outcome times the utility of that outcome.) But if the total number of consequences is not identified correctly, none of the probability assessments attached to any particular outcome can be valid. And this particular ill outcome should give pause to any utilitarian who sets a negative value on giving additional pleasure to a psychopath like Pedro. In that sense, Jim does make something worse happen by shooting the Indian. Certainly this unaccounted-for outcome should trouble Williams, since he sets a non-utilitarian, independent value on Jim's integrity, which has now been shattered "for nothing."

In fact there are four possible outcomes in the Jim and Pedro example: 1) Jim refuses to shoot the single Indian, and Pedro shoots all twenty; 2) Jim refuses to shoot the Indian, and Pedro, impressed by Jim's steadfastness and embarrassed by his own bloodthirstiness, changes his mind about shooting the others; 3) Jim shoots the Indian, and Pedro liberates the nineteen others; and 4) Jim shoots the Indian, and Pedro executes the others as well--the outcome I mention above. (There are of course other possible outcomes, such as the cavalry rolling up in the nick of time, but these really are extraneous to the agents concerned, Jim and Pedro.¹³) If outcome two occurs--and something similar has sometimes happened in cases of terrorists taking hostages¹⁴--the common response would be to admire Jim and rejoice in the happy outcome.

This casts doubts on whether Jim should abandon his integrity. But is the common response correct? Intuition allows the actual outcome to make a difference to how we assess the agent's choice; but intuition also holds that moral choice is somehow impervious to how things turn out. We commonly admire steadfastness, and this is the virtue which Jim would be said to display if the outcome turned out favourably, with Pedro backing down. But would we call Jim mule-headed or callous if Pedro did not back down? How can the character of the agent depend on the outcome? Thus the common response is incoherent insofar as it exemplifies the paradox of moral luck, and we have already seen that the paradox of moral luck threatens the notion of responsible agency. Yet the very virtue of steadfastness which would commonly be admired if Jim held out and Pedro backed down implies

that there is a wholeness to the agent which cannot be diminished by ill-luck in outcomes.

The impetus in both Williams and Nagel is this inconsistency in common sense, over the question of whether outcome luck can threaten responsible agency. Both begin with Kant as the source of the view that there can be a quintessential form of value, moral value, which is "unconditioned,"¹⁵ that is, free from external contingency. Both are also concerned at the outset to distinguish outcome luck from constitutive luck, that is, good fortune in having the "right" or the most praiseworthy inclinations, abilities, or temperament. This fourth variety of luck has some bearing on the question of whether good character can be destroyed by ill chance, and I shall discuss that in greater detail in chapter two, emphasising Nussbaum's concern with this question. It combines with the three other sorts of luck which I have already mentioned--luck in the problems which have to be faced, luck in antecedent circumstances, and luck in outcomes--to complete the typology which Nagel offers.¹⁶

Listing all these ways in which moral agency appears to be subject to chance leads to a heightened sense of the possible threat from moral luck, even though both authors mainly confine themselves to outcome luck. In Nagel and to a lesser extent in Williams, the purpose of glancing at other forms of luck than the outcome sort is to demonstrate how little might seem to be within the moral agent's control. Nagel presents this as apparently crippling not only for the Kantian enterprise, but also for the ordinary practice of attributing rightness and wrongness to actions:

[T]he broad range of external influences here identified seems on

close examination to undermine moral assessment as surely as does the narrower range of familiar excusing conditions. If the condition of control is consistently applied, it threatens to erode most of the moral assessments we find it natural to make.¹⁷

Again, "if one cannot be responsible for the consequences of one's acts due to factors beyond one's control, or for antecedents of one's acts that are properties of temperament not subject to one's will, or for the circumstances that pose one's moral choices, then how can one be responsible even for the stripped-down acts of the will itself, if they are the product of antecedent circumstances outside of the will's control?"¹⁸ This raises deep problems of identity as well as of responsibility. As Nagel puts it, "The inclusion of consequences in the conception of what we have done is an acknowledgement that we are parts of the world, but the paradoxical character of moral luck which emerges from this acknowledgement shows that we are unable to operate with such a view, for it leaves us with no one to be."¹⁹ Williams moves from this deep problem about identity as a responsible agent into two possible ways out of the paradox, on which I shall elaborate shortly.

One's history as an agent is a web in which anything that is the product of the will is surrounded and held up and partly formed by things that are not, in such a way that reflection can go only in one of two directions: either in the direction of saying that responsible agency is a fairly superficial concept, which has a limited use in harmonizing what happens, or else that it is not a superficial concept, but that it cannot ultimately be purified...²⁰

But are matters really this serious? Is "moral luck" is a genuine contradiction in terms, a true paradox? A useful borrowing here would be the recasting of the puzzle by Michael Zimmerman:²¹

1. A person P is morally responsible for an event e's occurring only if e's occurring was not a matter of luck.
2. No event is such that its occurring is not a matter of luck.

Therefore

3. No event is such that P is morally responsible for its occurring.

Nagel, according to Zimmerman, denies the conclusion but accepts both premises, believing the paradox to be genuine. Williams, he says, tends to accept the second premise while denying the first. But if this is so, Williams is not obliged to accept the paradox as genuine--still less to call for a complete overthrow of the Kantian ethical tradition because it is radically undermined by the paradox, or to doubt whether any doctrine of ethics can resolve the dichotomy. I suspect that Williams does accept the first premise as well. Certainly he is no less sceptical than Nagel, as Zimmerman's assessment would imply: in fact Williams presents himself as more sceptical than Nagel about the possibility of there being unconditioned moral value, untainted by luck.²²

Early in this chapter I suggested a focus on the weakness of premise two: although all events are a matter of luck, they are not all a matter of moral luck, nor are they all moral events. Zimmerman also proposes that we narrow the terms down. Thus, if premise one is reformulated as "P is morally responsible only if he was in restricted control of e," premise two cannot be maintained if it is in turn reformulated as "no event is such that anyone is ever in restricted control of it." To reverse the procedure, premise one becomes palpably false if refashioned as "P is morally responsible only if he was in unrestricted control of e", while two is true if revised to read "no event is such that anyone is ever in unrestricted control of it." Thus the distinction becomes that between restricted and unrestricted control.

Zimmerman believes that the paradox is false, and that recasting the terms dispels the moral luck problem. I do not think things are quite so simple, but I do believe that trying to apply limits to what agents are responsible for, with restricted and unrestricted control as background concepts, is a promising way to proceed. I shall explain why below.

Let us operate on the partially tested assumption, then, that the paradox of moral luck cannot simply be dispelled by recasting the terms. If matters are this serious, are they only grave for Kantians? Although Williams warns that "the aim of making morality immune to luck is bound to be disappointed,"²³ he also asserts that the paradoxes about moral luck are not merely some strange Kantian mania, that no notion of responsible agency can ignore them.

[Non-Kantians] may be disposed to think, so far as morality is concerned, that all that is in question is the pure Kantian conception, and that conception merely represents an obsessional exaggeration. But it is not merely that, nor is the Kantian attempt to escape luck an arbitrary enterprise. The attempt is so intimate to our notion of morality, in fact, that its failure may rather make us consider whether we should not give up that notion altogether.²⁴

To resolve the paradox of moral luck, we could absolve agents from responsibility for all matters beyond their ability to predict or control. But then we will wind up holding them responsible for very little. This is the strategy which Williams has pursued, and which is hinted at in this quotation. On some natural assumptions, it follows from Williams's assertion that only success can justify an agent's decision under uncertainty.

If he fails...then he did the wrong thing, not just in the sense in which that platitudinously follows, but in the sense that having done the wrong thing in those circumstances he has no basis for the

thought that he was justified in acting as he did. If he succeeds, he does have a basis for that thought.²⁵

If only success can justify a decision, but success is not certain at the time the agent makes a choice, there will turn out to be no basis but hindsight for judging whether an action was right or wrong. We do not generally think that agents are responsible for matters which they could only have known with hindsight. Therefore we will not be able to hold agents responsible for very many of their choices.

Alternatively, we could try to retain some notion of responsible agency but limit what agents are responsible for. This is a crude representation of Nagel's answer to the dilemma. My own view--extremely tentative at this point--is that Nagel is more likely to be right about how to resolve the paradox, although I do think that Williams's diagnosis of the paradox's gravity is correct. Moral luck is not only a problem for Kantians, as I argue further in chapter three. Nevertheless, I believe that Williams has despaired unnecessarily, although some commentators, such as Nussbaum, think that his despair has an attractive tragic-mindedness to it. It looks as if Williams is advising us to change our practice, to stop holding people responsible when they are not, and Nagel is telling us to come up with a different theory of agency, to which we should then try to make our practice conform. But actually Williams goes much further than this, almost to a refusal to have any further dealings with ethical theory at all--presumably because the notion of responsible agency, which is seen as crucial to any ethical theory, is undermined by the dependence of success on factors outside the agent's control.

Yet it is strange that Nagel should be more optimistic about retaining some notion of responsible agency, since he identifies more ways in which chance threatens that notion. The difference is that Williams retains a very broad notion of what counts as moral. In the end that turns out to be more crippling to the notion of ethical agency than an extensive catalogue of what counts as luck. In discussing a Gauguin-figure²⁶ whose decision to abandon his family responsibilities and stake all on aesthetic (if not commercial) success will only have been morally right if that success is forthcoming, Williams relies on the assumption that being a good or bad artist has some bearing on good or bad moral character. But suppose the twain simply do not meet? What if this is not an example of moral luck? As I said at the start, not all kinds of luck are moral.

Nagel agrees that this is not a moral luck example, even if failure as an artist undermines Gauguin's deepest sense of his own identity.

[Williams] points out that though success or failure cannot be predicted in advance, Gauguin's most basic retrospective feelings about the decision will be determined by the development of his talent. My disagreement with Williams is that his account fails to explain why such retrospective attitudes can be called moral. If success does not permit Gauguin to justify himself to others, but still determines his most basic feelings, that shows only that his most basic feelings need not be moral. It does not show that morality is subject to luck.²⁷

Williams counters this objection with what might appear a tautology: if Gauguin is injured on the way to Tahiti and never paints again, he cannot know whether or not his decision is justified. In that sense the injury is relevant to the moral decision, but only in establishing its parameters. It merely prevents him from knowing whether he was justified: "It is too

external for it to unjustify him, something which only his failure as a painter can do."²⁸

Although Williams retains a broad definition of what counts as moral, he does introduce an additional distinction to do with what happens if things go wrong.

From the perspective of consequences, the goods or benefits for the sake of which Gauguin's choice was made either materialise in some degree, or do not materialise. But it matters considerably...in what way the project fails to come off, if it fails. If Gauguin sustains some injury on the way to Tahiti which prevents his ever painting again, that certainly means that his decision (supposing it now to be irreversible) was for nothing, and indeed there is nothing in the outcome to set against the other people's loss. But that train of events does not provoke the thought in question, that after all he was wrong and unjustified. He does not, and never will, know whether he was wrong. What would prove him wrong in his project would not just be that it failed, but that he failed.²⁹

There are two points worth noting here. The more minor one is that Williams adds that utilitarianism misses something about this situation. This implies that utilitarianism could learn something by considering the paradox of moral luck, even though the paradox looks foreign to consequentialism at first glance. The major point, however, is Williams's distinction between the failure of Gauguin's project and Gauguin's failure as a moral agent.

In chapter five I shall draw on a similar dichotomy in discussing how medical professionals should feel if they choose one patient over another for receipt of a scarce medical resource, but the patient who receives the kidney or the expensive operation then dies. The doctors chose the "wrong" patient, but not necessarily on the basis of a faulty diagnosis of the two patients' chances of recovery. Since diagnoses are statistical judgments, a

certain proportion of them are bound to turn out "wrong". I present this as a moral luck case, but also suggest principles which would allow the case in which the doctors "get it wrong" to count as a project failure, occasioning regret, rather than a personal failure as a moral agent, which would dictate remorse. My handling of the two categories of regret and remorse is a little different from Williams's, however. In particular, Williams likes to use a term which I find rather unnecessary at least in the example of scarce medical resources: "agent-regret," which is felt by an agent about her past actions, rather than about external states which have gone wrong. Williams prefers "agent-regret" because he says it need not only relate to voluntary acts, as "remorse" conventionally does, but to any outcomes of a causal chain in which one has some role. (In chapter five I give an explanation of why I shall retain the more conventional terms, "remorse" and "regret".)

It seems, then, that not all project failures are personal moral failures. If Gauguin suffers the ill chance of injury on the way to Tahiti, his failure is neither personal nor moral. It is a failure of his project: and here I part company with Williams, agreeing with Nagel that no matter how deeply Gauguin feels about his failure, the intensity of that feeling does not itself make the failure a matter of ethics. (If it is an ethical decision, that is because Gauguin breached his responsibilities to his family.)

Now it might appear that we could get round these confusions by means of a scheme dividing factors necessary for personal or project success into

those extrinsic to the agent and those intrinsic to her, and Williams does try something like this:

The intrinsic luck in Gauguin's case concentrates itself on virtually the one question of whether he is a genuinely gifted painter who can succeed in doing genuinely valuable work. Not all the conditions of the project's coming off lie in him, obviously, since others' actions and refrainings provide many necessary conditions of its coming off--and that is an important locus of extrinsic luck. But the conditions of its coming off which are relevant to unjustification, the locus of intrinsic luck, largely lie in him--which is not to say, of course, that they depend on his will, though some may. This rough coincidence of two distinctions is a feature of this case. But in others, the locus of intrinsic luck... may lie partly outside the agent, and that is an important, and indeed the more typical case.³⁰

The example which Williams uses to illustrate this "more typical" and thus perhaps more important case is a similarly schematized version of Anna Karenina. He takes Anna's suicide to be an admission that she has failed, not merely that her project has failed through some extrinsic factor--for example, if Vronsky had been killed in an accident. "What she did, she now finds insupportable, because she could have been justified only by the life she hoped for, and those hopes were not just negated, but refuted, by what happened."³¹

If we accept Williams's interpretation of the Anna Karenina story, grave paradoxes certainly ensue. Anna could not have known in advance that her affair would end badly, but she could only have been justified by its turning out well. What basis could she then have had for her decision about whether to leave her husband? This looks at first to be the problem about hindsight and responsibility which I identified before.

But actually to me this example is mainly about what counts as an ethical choice and what agents are responsible for. The case of Anna

Karenina cannot be the universal example which Williams thinks it is, because in Anna's society only women had to make such a choice. Admittedly, Anna is a schematised figure, as was Gauguin, but this does not get round the problem of non-universalisability. There is no moral luck in this example because there is no genuine moral choice or responsibility. Here I would adopt a Kantian position, of saying that the genuine moral life must, by its nature, be equally open to all, and its dictates equally applicable to all. Because men were free of the strictures which required Anna to make a choice, because they operated under a much less restrictive system of responsibilities than did women in Anna's society, the system of "virtue" which obtained in Anna's time cannot count as a genuinely moral system, nor can her failure to live up to it count as a moral failure.

I want to suggest that perhaps Anna should be viewed as a victim rather than as an agent making a moral choice. This is not to accept Williams's very broad assertion that no one can definitely be said to make moral choices, because the entire existence of such a thing as ethical choice is in doubt. I also want briefly to ask how Anna could have succeeded, given the systematic way in which the odds were stacked against her, and to suggest that the fact that she had to make a decision at all is evidence of that sort of discrimination.

Anna had to decide between husband/son and lover. Her brother Oblonsky was allowed, under the double standard, to keep both his mistresses and his home life. "Although Oblonsky was entirely in the wrong as regards his wife, as he himself admitted, almost everyone in the house, even the nurse,

[his wife's] best friend, was on his side."³² With a fine and deliberate irony, the novel begins with Anna persuading her sister-in-law to forgive Oblonsky for his latest dalliance. Williams seems to want to argue that Anna's affair was wrong from the start because it ended badly. Her brother's affairs don't end badly: are they any less wrong?

Williams does hone his analysis by noting that once Anna left her husband, the relationship with Vronsky would have to bear too much weight. He calls her failure to foresee this an intrinsic matter. Certainly the novel's description of the lovers' scenes confirms that the relationship is overloaded; but that is only because society will not allow Anna to retain the marital relationship as well.³³ Although Williams calls this failure to forecast the extra burden on the relationship "a truth not only about society but about her and Vronsky,"³⁴ I cannot see that the occasion of the moral decision would have arisen but for society's extrinsic laws and mores.

Still less do I agree with Judith Andre's contention that we would be right to praise Gauguin and condemn Anna Karenina because "the person who can correctly assess his or her chances of success is better-formed than the person who cannot."³⁵ This confuses acting ethically with being a successful technician. But even if there were a moral virtue in being well-calibrated, the point is not proved by the contrast of Gauguin with Anna Karenina.

First, we do not know whether Gauguin judged the probability of success accurately or inaccurately; we only know that he succeeded. Old Laminitis, the 100-1 shot whom I decide to back in the 2:30 at Towcester because I think his name is a good joke, may actually romp home first against the odds, but that says nothing about my knowledge of form. Second, although it is arguable whether Anna has the stature of a tragic heroine, she does face a tragic dilemma in which there is no chance of unmitigated "success". Provided that we accept her love for Vronsky as a fact of the situation rather than as something she could do something about, both her possible choices entail unavoidable penalties. Her lot will be very different from her brother's if she, like him, tries to have it both ways. If she tries to carry on her affair under her husband's nose, he has the legal right and the economic clout to make her life very uncomfortable. If she runs off with Vronsky, society's sanctions will be the predominant ones. Of course if Gauguin fails to measure up to his own artistic standards, he faces self-condemnation, but there is no external social penalty, as there is for both Anna's choices. Either way Anna loses. The probabilities of full-fledged success attaching to the two outcomes are irrelevant because they are equal, at zero; she has not in fact demonstrated faulty calibration by opting to run off with Vronsky.

I have criticised both of Williams's examples, as not being specifically about moral luck. In the process I have begun illustrating how limiting what counts as moral might enable us to retain a notion of morally responsible agency, of which Williams despairs. But perhaps this is only because he stakes so much on these two examples, which I consider faulty.

Then what sorts of examples would constitute genuine moral luck cases? It would be possible to stipulate that good examples must involve agents who stake their moral success--not their worth as artists, their happiness, or any other value--in advance on outcome luck. I believe that there are many such examples to be found in medical ethics.

Take the case of a hospital ethics committee which approves random clinical trials of a new drug on bowel cancer patients without their consent. A patient with a prognosis of several years of active life dies two weeks later because her bone marrow becomes irreversibly depressed after administration of the drug.³⁶ Let us suppose that no negligence is present, that the drug was administered correctly but turned out to have side-effects which no one could foresee. The experiment has gone wildly wrong, for reasons which are not the committee's fault. (For the time being, I am equating two matters I shall separate out in chapter three, prediction and control.) But should we still hold the committee members responsible for the patient's death? In chapter seven I look at several such cases, in the area of informed consent.

Perhaps because I am not convinced about Williams's two examples, I am not entirely persuaded by his fear of the "final destruction" which occurs when Kantian strictness is united with a utilitarian doctrine of negative responsibility.

There is, at the end of that, no life of one's own, except perhaps for some small area, hygienically allotted, of meaningless privacy. Because that is a genuine pathology of the moral life, the limitation of the moral is itself something morally important.³⁷

Here I do agree with Williams, but I think he is now pointing towards the

second path out of the paradox, rather than towards his generally preferred first solution. That is, rather than saying that we can have no meaningful conception of responsible agency, he is indicating that we can have such a conception if we limit what agents are responsible for. This is close to my own hypothesis, and I shall discuss this question further in chapters three and four.

However, it is the first, rather deliberately tragic-minded, course which Williams has pursued in later work, notably in Ethics and the Limits of Philosophy. In this book he rejects the Kantian claim even more emphatically: "[T]he idea of a value that lies beyond all luck is an illusion."³⁸ Part of Williams's motivation is the attempt to avoid circularity: "the standpoint from which pure moral value has its value is...only that of morality itself. It can hope to transcend luck only by turning in on itself."³⁹ Now this raises an interesting line of argument which I pursue in chapter five, to do with drawing luck's fire by recognising it openly. But I am sceptical about Williams's subsequent solution to the problem about unconditioned value. That is, roughly, that we could substitute for Kantian morality something approximating to an Aristotelian conception of the good whole life for a man.

In thinking that a culture could agree on values, even if the values were not given the full weight of ethical norms, Williams is making a pluralist assumption of consensus. That is, he merely assumes that society is not radically divided by structural imbalances of gender, race or class. (In chapter five I criticise in similar terms one schematic analysis of

"society's" criteria for who shall receive scarce medical resources.) In any case, it is through reviving this conception from the ancients--whether or not the device succeeds--that Williams tries to mitigate the paradox of moral luck. This leads us neatly into the concerns of chapter two.

FOOTNOTES

1. Kant, Fundamental Principles of the Metaphysic of Morals, first section, third paragraph, quoted in Thomas Nagel, "Moral Luck," in Mortal Questions (Cambridge: Cambridge University Press, 1979), p. 24.
2. Martha Craven Nussbaum, The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy (Cambridge: Cambridge University Press, 1986). Two lengthy critical reviews are sceptical of Nussbaum's claim, regarding it as anachronistic: John M. Cooper's article in Philosophical Review, v. XCVII, no. 4 (October 1988), 543-563, and Nicholas P. White, "Rational Self-Sufficiency and Greek Ethics," in Ethics, v. 99, no. 1 (October 1988), 136-146. These criticisms, along with Nussbaum's work, will be discussed in greater detail in chapter two.
3. Proc. Aris. Soc., supplementary volume L, 115-135 (Williams) and the reply by Nagel in the same volume.
4. Judith Andre, "Nagel, Williams and Moral Luck," Analysis, v. 43 (1983), 202-207; Michael J. Zimmerman, "Luck and Moral Responsibility," Ethics, v. 97, no. 2 (January 1987), 374-386.
5. Claudia Card, "Women's Voices and Ethical Ideals: Must We Mean What We

Say?", review of Women and Moral Theory, ed. Eva Feder Kittay and Diana T. Meyers (Totowa, New Jersey: Rowman and Littlefield, 1987), in Ethics, v. 99, no. 1 (October 1988), 125-135.

6. Bernard Williams, Ethics and the Limits of Philosophy (London: Fontana/Collins, 1985).

7. This seems the logical location for a discussion of Nussbaum, since she herself begins by castigating Kant--even though most of the book is rooted in classical philosophy and literature. One critic, Nicholas P. White, has complained that Kant becomes a stand-in for Plato, the real target of Nussbaum's dislike. If this is so, it would perhaps have been better for me to have considered Nussbaum's important work in a separate chapter on the classical philosophers. But John M. Cooper also complains that there is little evidence that Plato or Aristotle was actually motivated by worries about the contradictions inherent in moral luck, and if this is true there can hardly be a separate chapter concerning the moral luck debate in the ancients.

8. Ibid., p. 140, for the germ of this example.

9. Nagel, p. 28.

10. From Williams, "A Critique of Utilitarianism," in J.J.C. Smart and Williams, Utilitarianism For and Against (Cambridge: Cambridge University Press, 1973). Williams advances the example in a discussion of negative responsibility, and further discussion of it has often been in this light: see John Harris, Violence and Responsibility (London: Routledge Kegan Paul, 1980), p. 110 ff. I do not intend to deal with negative responsibility, only with the decision-theoretic question of whether Williams has given us all the possible outcomes. For simplicity's sake, I use Harris's recast

version of the example: in the original Pedro is the captain's helper.

11. Ibid., p. 117.

12. It may appear contradictory to call this outcome a mere matter of luck, if luck means unforeseen contingency, since I also imply that Jim, and Williams, should take the possibility of it into account. In fact I do not define luck as unforeseen contingency, merely as contingency. In any case, although the possibility of Pedro going back on his word is foreseeable, whether it will occur this time is unpredictable insofar as statistical independence enters in: just as the probability of a coin coming up heads is definitely .50 but the occurrence of heads on the next throw entirely uncertain. This point will be dealt with at greater length in chapters three and four.

13. In considering both Jim and Pedro as agents, I follow Williams: "While the deaths, and the killing, may be the outcome of Jim's refusal, it is misleading...to leave Pedro out of the picture in his essential role of one who has intentions and projects, projects for realising which Jim's refusal would leave an opportunity." (Williams, "A Critique of Utilitarianism," p. 109.)

14. Williams has dismissed the possibility of there being other outcomes than one and three in the Jim and Pedro case, regarding these two as certain beyond reasonable doubt. But in several sieges, such as those at Spaghetti House and the Polish embassy in Bern, the terrorists did in fact release the hostages even though the authorities refused to cede to their demands. It seems that Williams's impatience is directed at a general political argument which suggests that "the efficacy of the detestable action is more doubtful than the example supposes" ("Utilitarianism and

Moral Self-Indulgence, in Moral Luck, p. 43). "This is a line often taken by those defending an absolutist position in cases of detestable actions extorted by threats made by hijackers and so forth, to the effect that the very character of the threat shows that one has reason to doubt the efficacy of giving in to it. Why should one expect such threateners to keep their promises anyway? As a general line of argument, this seems, bluntly, a cop-out." (original emphasis). Now to me, the cop-out is attempting to write off possible outcomes which the moral agent ought to consider, assuming that she is considering consequences at all. However, I do not commit myself to a general absolutist line such as Williams castigates. It may be that our disagreement is only over whether the Jim and Pedro example is one of those "cases in which it is a reasonable bet that nothing is to be gained by giving in to threats," whose existence he does recognise (also p. 43). Even more narrowly, all I want to suggest is that it ought to be permissible to consider whether the Jim and Pedro example is such a case.

15. Williams, "Moral Luck," p. 20, attributing the term "unconditioned" to Kant. Note the consistency of defining chance merely as contingency, rather than as that of whose likelihood nothing can be known (see footnote 12). This is also consistent with decision theory, much of whose purpose is to minimise our ignorance about the workings of chance.

16. Nagel, "Moral Luck", p. 28.

17. Ibid., p. 26.

18. Ibid., p. 35.

19. Ibid., p. 38.

20. Williams, "Moral Luck", p. 29.

21. Zimmerman, "Luck and Moral Responsibility", p. 374.

22. Williams, "Moral Luck", p. 36, fn. 11.
23. Ibid., p. 21.
24. Ibid., p. 22.
25. Ibid., p. 23.
26. Williams refuses actually to identify him with the historical Gauguin: he is a hypothetical example ("Moral Luck", p. 22).
27. Nagel, "Moral Luck", p. 28, fn. 3.
28. Williams, "Moral Luck", p. 25.
29. Ibid., p. 25.
30. Ibid., p. 26.
31. Ibid., p. 27.
32. L.N. Tolstoy, Anna Karenin, tr. Rosemary Edmonds (Baltimore, Maryland: Penguin Books, 1954), p. 17.
33. Karenin refuses to grant Anna a divorce, so that the formal marriage persists, but not the marital relationship.
34. Williams, "Moral Luck", p. 26.
35. Andre, "Nagel, Williams and Moral Luck", p. 206.
36. This is an actual case, the only difference being that an element of negligence was also present: the effects of the drug were monitored carelessly. The trial was in fact approved by eleven hospital ethics committees, not one. See Carolyn Faulder, Whose Body Is It? The Troubling Issue of Informed Consent (London: Virago, 1985), p. 9.
37. Williams, "Moral Luck", p. 38.
38. Williams, Ethics and Limits, p. 196.
39. Ibid., p. 195.

Chapter Two

Risk and Kant

"If there remained only the good will...it would sparkle like a jewel in its own right, as something that had its full worth in itself." To Kant the good will is implicitly beautiful; but to Martha Craven Nussbaum, the good cannot be "beautifully human" if it is distanced from chance.¹ Its fragility is what makes human endeavour admirable and touching, and the Kantian attempt to create a risk-free ethics, as she interprets it, impoverishes the branch of human endeavour called acting well. (This is particularly serious if the moral life is the highest branch of human endeavour, taking precedence over the good life, and I shall return to this question shortly.) Nevertheless, Nussbaum recognises that there is no reason to aim at moral excellence if it cannot be called one's own, if it is inherently vulnerable to chance. "The question of the human good" thus becomes: "How can it be reliably good and still be beautifully human?"²

"Human excellence grows like a vine tree, fed by the green dew, raised up, among wise men and just, to the liquid sky." Taking her metaphor for virtue from Pindar, Nussbaum implicitly rejects Kant's jewel parallel:

[T]he poetic image...suggests that part of the peculiar beauty of human excellence just is its vulnerability. The tenderness of a plant is not the dazzling hardness of a gem. There seem to be two, and perhaps two incompatible, kinds of value here.³

Still, the possession of reason is also part of human excellence: part of

this particular plant's genetic blueprint is to develop the cutting hardness of a diamond. In some sense, according to Nussbaum's interpretation of the Greeks, we are less vulnerable to chance than is the rest of creation, because through reason we exercise some measure of control. Presumably this would be mainly through partial foreknowledge of what is going to occur, although through rational persuasion we might sometimes control the actions of others, and through the application of scientific reason we might develop some control over nature.

Notwithstanding, reason does not give us full control. The fragility of goodness--in Nussbaum's fine phrase--is confronted openly in Aristotelian philosophy and the Greek tragic poets, she says. But according to Nussbaum, Kantianism refuses to recognise that the goodness or badness of character and the rightness or wrongness of actions simply do hinge on chance, on matters beyond the agent's control. This wilful blindness, combined with what she views as the dominance of Kantian thought throughout the modern period,⁴ has persuaded us that somehow tragic conflicts can be mitigated, if not avoided. We delude ourselves into thinking that we can still retain a sense of agent-responsibility even when agents are faced with two equally evil alternatives--as Agamemnon was at Aulis, when, on Zeus's arbitrary instructions, his only choice lay between allowing the plague to continue among the becalmed Greek ships and sacrificing his daughter. For Agamemnon to have been "morally lucky", in the Greek view, would have been to have escaped this quandary altogether; but once it was visited upon him, it was impossible for him to retain his moral integrity. No Kantian notion of a good will or pure intentions could have saved him

from this ill-luck. (It should be noted, however, that this is what Nagel calls luck in the decisions which have to be faced, not outcome luck--the main focus in Kant and in this dissertation. Nussbaum does not always separate out the various forms of luck, a criticism to which I shall return later.)

Because of our Kantian presuppositions, Nussbaum says, we are prone to distance ourselves from the Greek approach, to deny that some dilemmas simply are irreparably tragic, to assert "that Greek thinkers held the false and primitive view that moral value is vulnerable to luck."⁵ To Nussbaum, this is a catastrophe for ethics, almost as grave as the mythical one with which MacIntyre begins After Virtue.⁶ Moreover, the Kantian, in Nussbaum's presentation, makes a watertight distinction between ethical value and other sorts of value, as she says the Greeks did not. Not only does the good will shine like a jewel in its own right, having its full worth in itself; its worth is also greater than that of anything else in the human sphere. This would lead the Kantian to a second, related conclusion about why Greek ethical thought was primitive: it did not even bother to distinguish between the virtuous life and the good human life.

When the truth of these Kantian beliefs, and the importance of the Kantian distinction between moral and non-moral value, are taken as the starting-point for inquiry into Greek views of these matters, the Greeks do not, then, fare well. There appears to be something peculiar about the way they agonize about contingency, lamenting an insoluble practical conflict and the regret it brings in its wake...It is as if they were in difficulties because they had not discovered what Kant discovered, did not know what we Kantians all know.

Now the point for my purposes is not whether the Greeks have been hard done by; it is whether this is an accurate picture of Kantianism in

relation to moral luck. Basically, Nussbaum charges Kant with ignoring the dilemma of moral luck, with turning away from the myriad ways in which chance affects how things turn out and focusing only on the purity of the will. Recall that the paradox of moral luck arises from the tension between (1) the way in which responsibility depends on some things outside moral agents' control and (2) our unwillingness to blame people for things beyond their control. Nussbaum accuses Kant of ignoring the first pole of the dichotomy. But this conflicts with the rough sketch of Kantianism which I presented in chapter one. There I said that the paradox derives from a Kantian outlook, from asking Kantian questions.

To charge Kant with ignoring reality also overlooks his remark in the Lectures on Ethics that a deathbed repentance is not a genuine act of conscience because it has no chance of being carried into practice.⁸ Kant wants conscience to have an effect on the world; does he also want the world have some impact on the integrity of conscience? How would this be compatible with his assertion that the good will is untouched by how things turn out? In this chapter I want to defend Kant against the charge of ignoring the dilemma of moral luck, and to ask whether a Kantian solution to the paradox looks like the most promising one after all. Because Kant proposes a solution does not mean that he ignores the problem.

But I shall deal first with Nussbaum's secondary charge against Kant, that he impoverishes the good life by narrowing it to the moral life. Thomas Nagel identifies more possible positions on this question than Nussbaum allows. In criticising the use of a similar dichotomy by Bernard

Williams, Nagel distinguishes five rather than two possible relationships between moral and other sorts of value:⁹

1. The moral life is defined in terms of the good life, the position attributed by Nagel to Aristotle. The two sets of values are not seen as equivalent, but the content of the ethical is dictated by background conditions for the good life, particularly the good life as appropriate for a certain social station.

2. The good life is defined in terms of the moral life. Both the first two positions deny that there is a conflict between the good and moral lives: Nagel reads the Republic as an elaborate and "heroic" justification of this denial, and Plato as demonstrating empirically what Kant can only postulate, "that moral virtue forms an indispensable part of the good for each person."¹⁰ The next two positions admit that there is a disparity between the good and the moral lives.

3. The good life overrides the moral life--the view taken by Nietzsche, Thrasymachus in the Republic, and Philippa Foot in her later work, Nagel thinks, and perhaps also by Gorgias, Polus and Callicles in the Gorgias, I would add.

4. The moral life overrides the good life. Although Nussbaum identifies this argument as quintessentially Kantian, Nagel regards it as typical of utilitarianism and rights theory as well. The link between the two charges levelled by Nussbaum at Kantianism is highlighted by Nagel's

observation that position one defines the good life as what is best for a particular individual, distinguished from others by sex, age, social standing and other personal characteristics. Position four makes no such distinctions; it is concerned with the totality of humans, Nagel asserts--and, I would add, with the statistical totality of events, in some consequentialist formulations. "Any coincidence between this [the good for all humanity] and what is best [for the individual] will be a matter of luck, or political and social arrangement."¹⁰ The importance of "political and social arrangement" as an adjunct to luck is that it helps to explain cases like Anna Karenina's, which Williams misconstrued, I argued in chapter one, because he confused factors intrinsic to Anna's self-worth with those external social biases which doomed her project.

On this matter, I think Kant is more coherent. Unlike the ancients (particularly Aristotle), he refused to accept a system of virtues in which the ability to attain praiseworthy behaviour depended crucially on station in society or on other chance factors of birth. This is the relevance of his reference to the "niggardly provision of a stepmotherly nature". A modern corollary of this Kantian view is Rawls' insistence that intelligence is not a matter of desert or merit, but is conferred by good fortune; therefore rewards should not flow from it, under a fairness-orientated system of justice.¹² Another modern application of the Kantian principle that the virtuous life must be equally accessible to all might be a feminist condemnation of the classical system of virtues, in which the ill-chance of being born female limited a woman to a single, confining virtue, sophrosyne or self-restraint, and barred her from the more highly

rated heroic qualities. But this would be an egalitarian reworking: in Kant this requirement of equal accessibility to the virtuous life flows from the demands of the ethical enterprise itself, not from any egalitarian impulse.

5. Neither the good life nor the moral life consistently overrides the other.

Nagel considers position five attractive but comes round to the view that four is probably correct. Clearly he disagrees with both Williams and Nussbaum, with their mutual preference for position one--and he might well accuse them of failing to distinguish one from three. What matters here is not so much his reason for approving the fourth view as his reminder that there are more than two possible relationships between the good and the moral lives. If so, the second of Nussbaum's charges against Kant is an oversimplification. Moreover, it is inaccurate, in Nagel's typology, to conflate Kant and Plato. To borrow a term from Williams, Plato's project is very different from Kant's, as Nicholas White has asserted in a critical review of Nussbaum.¹³ White thinks that Williams is tarred with the same brush, but not from head to foot: Nussbaum is writing a major study of the Greeks, whereas Aristotle and (even more so) Plato are only peripheral to Williams's work. However, as I pointed out before, my concern in this chapter is not the accuracy of Nussbaum's reading of Plato or Aristotle: only her faithfulness to Kant. But if she is using Kant as a stalking horse for Plato, there are grave reasons to be suspicious about the accuracy of her picture of the former.

What about Nussbaum's first charge against Kant--wilful negligence of the moral luck paradox? For my purposes, this is the more important accusation, to which most of the rest of this chapter will be dedicated. But there is a parallel between Nussbaum's failure to consider all five possible couplings of the good and moral lives and her failure to distinguish among the four sorts of luck which Nagel identified in "Moral Luck" as having a bearing on ethical worth. (See chapter one.) White, who is extremely (and in the end, I think, overly) dismissive of Nussbaum's work as ill-formulated, agrees that she is guilty on both these counts. He believes that in the Republic Plato advocates philosophical thought because it is good in itself, but not, as Nussbaum claims, because it is impervious to the vagaries of chance, because it is "unlosable." The ethically right does have its full worth in itself, but "all of that is compatible with its being as fragile as you like."¹⁴ Giving money to buy food for the hungry is always right, but whether the food actually gets through and the good deed has any effect is eminently vulnerable to blocked roads, floods, civil wars and flat tyres. Nussbaum, White charges, lumps together all sorts of luck as instability, but as I argued in the first chapter, not all luck is moral. And as John M. Cooper has pointed out, the sort of luck from which Plato, at least, wants to liberate us may not have much to do with the kind which concerns Williams.¹⁵

In short, Nussbaum may be prone to oversimplification, and this failing could affect the accuracy of her portrayal of Kant. White particularly censures her for what he views as an unfair attack on Kantianism as if it were a monolithic whole. It may be indicative that

Nussbaum makes no attempt to differentiate between Kant's various ethical works, and in particular that she makes no reference in her bibliography to the Lectures on Ethics. The translator of the Lectures, Lewis White Beck, argues that it is in them, and also in the Metaphysics of Morals, that we glimpse the Kant whose lively conversation and striking way with words are said to have charmed audiences of students and enlivened drawing-room repartee. In the Grundlegung and the Critique of Practical Reason, Beck says,

[e]verything anthropological and narrative or anecdotal, everything that would make perspicuous the relation of philosophical morals to the conduct of life, is there apparently sacrificed for an abstract intellectual articulation. It is no wonder, then, that Kantian ethics has since appeared to be forbidding in its intellectualism; to be rationalistic at the expense of emotion, habit, and institutions in the make-up of the good life; to have sacrificed all the graces for a few of the virtues...[Only in the Metaphysics of Morals and the Lectures do we see what Kant never forgot, but what he expected his readers to remember even when he was talking of other things--viz., that the good life is more than mechanical obedience to the categorical imperative, that right action requires more than right thinking, and that man is more than a thinking machine.]¹⁶

White also emphasises that we must keep Kant's view that there were no genuinely tragic choices separate from his position on moral luck.

Although Kant seems to have believed that there are no genuine conflicts of duty, he was certainly not forced to that belief by his doctrine that the only genuine value is independent of contingencies. For he could have maintained that doctrine and still held that when a person is confronted with a genuine conflict of moral incommensurables, the goodness of his or her will is unaffected by making one choice or the other, so long as the decision is made in full consciousness of the existence of both obligations (which is not far from what Nussbaum herself seems to maintain.)¹⁷

This chapter will now concentrate on Kant's views about risk and moral luck, but will include a short section on tragic choices at the end. My main argument here is that Kant is not so much concerned with freedom from moral risk as with freedom of the will, and that the ultimately

undetermined action of the will presents a sort of risk in itself--one which Kant in fact recognises. I want to show that Nussbaum could accept Kant if she had read him correctly. More broadly, this entails dismissing Nussbaum's criticism that Kant has simply ignored the way in which luck impacts on ethical choice.

Nussbaum seems to accept the commonly held view that Kantianism sets itself up as a risk-free system of ethics. Here she and Nagel would agree:

In Kant a course of action that would be condemned if it had a bad outcome cannot be vindicated if by luck it turns out well. There cannot be moral risk.¹⁸

To Nussbaum, a risk-free system of ethics is not worth having. Somehow the moral enterprise hardly seems worthwhile if it does not recognise how deeply the rest of life is imbued with chance. This risk is so profound, she argues, as to threaten even the good will. Using the example of Euripides's *Hecuba*, she asserts that even good character may decay through chance reversals. This possibility is deeply repugnant to Kantianism, she says: in terms of "a moral philosophy that speaks of the incorruptability of the good will, sharply distinguishing the sphere of contingent happenings from the domain of the moral personality, itself purely safe against the 'accidents of step-motherly nature'...this play tells dangerous lies."¹⁹

But the will is not the same as character in Kant; it is practical reason. A Kantian could perfectly well accept that good character is partly a matter of luck, just as intelligence is to Rawls. Both are normally considered achievements or personal virtues, but could equally well be

partly matters of chance. Nagel does treat character as a fourth type of luck, the constitutive variety..

In any case, this is a different criticism from the one about whether the good life should be confined to the moral life, as I have said: that one had to do with the narrowness of the Kantian vision. That insularity Nussbaum presents as itself a moral fault, in her discussion of Antigone, whose heroism is lessened by "a ruthless simplification of the world of value which effectively eliminates conflicting obligations...[S]he can be blamed for refusal of vision."²⁰ But that is a common enough criticism of Kant,²¹ and a more limited one. It does not question the worth of the entire Kantian ethical enterprise as does the assertion that a risk-free morality is not worth having. But in turn, the bigger charge is only a threat if Kant is in fact concerned to "risk-proof" ethics. I shall argue that this is not so: in Kant the moral enterprise is saturated with risk.

One obvious way in which Kantian ethics deliberately opens itself to the full vagaries of chance is that it refuses to excuse morally wrong decisions which happen to turn out well. Nagel notes this, but does not present it as an assumption of extra risk. Of course there are many cases in which bad judgement is saved by good luck. I argued in my first chapter that Gauguin might be one such example, although Judith Andre seems to think that his judgement must have been good because the decision turned out well. Here Nussbaum agrees with me, but appears not to realise that by so doing she also agrees in part with Kant.

We must avoid from the beginning a confusion between the assessment of the decision and the assessment of the deliberations that led to the decision. It is perfectly possible for a person to have reached

the better overall decision through a deliberative process that neglects certain valid claims; the decision will still, then, be correct--but not for the right reasons, and almost, as it were, by accident.²²

But Nussbaum regards the decision as still correct in some sense, even though correct by chance. The Kantian standard is stricter, and also more stringent than everyday attitudes towards blame and praise, however similar to common sense it may appear. No matter whether the action turns out well or badly: if we disobey the moral law "we feel its power even when we are most defying it...[T]he moral law itself, unlike any motive of desire, propels [us] onward to destruction."²³

This looks at first as if it could have very little to do with freedom in the moral agent. Nevertheless, the moral law, the categorical imperative, obtains its legitimacy and force from agent-freedom. It alone is unconditioned, to use the terminology which first appeared in relation to Kant in chapter one.

The categorical (unconditioned) imperative views the action as objectively necessary and necessitates the agent to it immediately, by the mere thought of the action itself...and not mediately, by the thought of an end to be attained by the action...All other imperatives...are, one and all, conditioned...The ground of the possibility of categorical imperatives is this: that they are based simply on the freedom of the power of choice, not on any other characteristic of choice (by which it can be subjected to a purpose).²⁴

It is precisely our freedom which enables the moral law to exist, in a two-way relationship. "...[T]hough freedom is certainly the ratio essendi of the moral law [the reason the moral law can exist], the latter is the ratio cognoscendi of freedom [the reason we know that we are free]".²⁵

Now there is obviously a link between my action being free and its being genuinely mine.²⁶ And we have seen that Nussbaum--rightly, I think--denies that there is any reason for me to aim at moral excellence if it is not truly mine. So there will be a connection between my action being free and its being worthwhile. This is why I shall argue that Nussbaum is wrong to view the Kantian moral enterprise as unworthy of us. In the Lectures on Ethics Kant actually asserts that someone who disobeys the moral law about truth-telling loses his selfhood: "A liar, even though by his lies he does no harm to anyone, yet...he throws away his personality."²⁷

Since I do not intend to devote a chapter to the classical philosophers--there being some doubt as to whether it is anachronistic to treat them in terms of moral luck--I shall not consider at any length Nussbaum's preference for Aristotelianism as truer to what it means to be human. But in passing, it is useful to point out that if an action is worthwhile because it is freely chosen, there must be serious doubts about whether Aristotelianism really is more attractive. Freedom to choose one's moral position would have been foreign to the Greeks, MacIntyre argues:

A man who tried to withdraw himself from his given position ... would be engaged in the enterprise of trying to make himself disappear...For freedom of choice of values would from the standpoint of a tradition ultimately rooted in heroic societies appear more like the freedom of ghosts--of those whose human substance approached vanishing point--than of men.²⁸

There is another reason why it is incorrect to depict Kant as attempting to provide a risk-free ethics. In Kant the risk-free course, doing only what is required, carries no chance of blame, but also no

possible gain in praise. "If I do exactly what is required of me, the consequences are neither my fault, nor are they to my credit."²⁹ It is the freedom of the agent which allows us to attribute responsibility. "In a word, the key to the imputation of responsibility for consequences is freedom."³⁰ Kant goes on with the familiar dictum that the good will is not good because of its consequences; but neither is it good because it is determined by duty. The good will can never claim to be simply following someone else's orders. Orders can be given either by an external authority, by law, or by an internal authority, practical reason. It is only the principles dictated by the latter which qualifies as ethical pronouncements in Kant. An external law may have the same content as an ethical command, but we cannot claim to have behaved well if we obey the law out of simple compulsion.

Duty does not subordinate the will, though it determines it temporally. Instead, duty is central in Kant because it is through the immediate perception of duty that the will realises its unconditioned freedom. Kant simultaneously presents duty as universal and also as entirely contingent upon each individual freely giving it to himself or herself. "Kant continued to regard the paradox of human freedom as unavoidable: we could never solve it through theoretical reason, while practical reason assured us only that it has a solution."³¹ Freedom is also the distinguishing mark of an ethical choice as opposed to an action taken in obedience to the external law.

In the moral sphere compulsion has no place; no one can compel us to acts of kindness or charity. Thus moral omissions and their consequences can never be imputed, but legal omissions can. Conversely, moral acts of commission with their consequences can be imputed, but legal acts of commission cannot, since they are

obligatory acts...I do not impute in demeritum the consequences of an action which a person is obliged to perform, because in such circumstances he has ceased to be free. He is responsible for the factum in itself but not for the illegitimacy of the factum.³²

This is another way of limiting what agents are responsible for while retaining a notion of agency. By this standard, Agamemnon is of course responsible for killing his daughter, but not for the killing being an evil deed. At first this looks puzzling: how can killing ever be anything but an evil deed? But the point is not whether it could ever have any other moral content; it is that Agamemnon is responsible only for the bare action, not the action's content. Zeus is responsible for that, presumably, if gods can be held responsible in the ordinary mortal way. But because Agamemnon is not to blame for the deed's illegitimacy, he should not try to justify it. Perhaps, indeed, he is only guilty if he tries to justify the action as being somehow legitimate. (This, of course, is exactly what he does try to do, as I shall discuss at the end of this chapter, and this is why Nussbaum blames him--for what may be seen as surprisingly Kantian reasons.) Similarly, and more broadly, we can be held responsible for setting events rolling through actions which we could control, but not necessarily for all the consequences of those actions.

Whatever appertains to freedom can be imputed to us, whether it arises directly through our freedom, or is derived indirectly from it. A drunken man cannot be held responsible for his drunken acts; he can, however, for his drunkenness. The causes which make it impossible to impute responsibility to a person for his actions, may themselves be imputable to him in a lower degree.³³

This will turn out to be an important way of limiting agent responsibility in the medical ethics examples in chapters five and seven.

It means that not only is there a connection between my action being free and its being truly mine; there is also a link between my action being free and its being invulnerable to moral luck. What can be said to belong to the drunk man and what he can be held responsible for, regardless of whether or not his drunkenness causes any actual harm to others, is his decision to drink heavily in the first place.

This is also apposite to an example given by Nagel in "Moral Luck", that of a drunk driver who swerves dangerously onto the pavement. Nagel does not actually take a fully Kantian line here: he asserts that the driver is morally lucky if there are no pedestrians on the pavement, and morally unlucky if there are. "If there were, he would be to blame for their deaths, and would probably be prosecuted for manslaughter."³⁴ But this is to confuse moral and legal responsibility in the way Kant warns us against. Intuitively we do think that the driver should say to himself, "I could have killed someone," and stop combining alcohol and driving. We do blame him whether or not an accident occurs, presuming that we get to hear of his conduct, and whether or not he is punished by the force of the law. This Kantian way out of the paradox of moral luck, which common sense has created, is itself consistent with common sense. Kant would regard this as a considerable virtue, I think. In dismissing the criticism that the Grundlegung contained "no new principle of morality in it but only a new formula," Kant argued, "Who would want to introduce a new principle of morality and, as it were, be its inventor, as if the world had hitherto been ignorant of what duty is or had been thoroughly wrong about it?"³⁵

In taking the "dry stick" view of Kant, as someone who tried to seal himself and his philosophy off from the world of chance and the emotional life, Nussbaum--and others--may simply have been influenced by the fabled regularity of Kant's daily routine, in which freedom appeared to play little part. Or perhaps she is over-emphasising the negative conception of freedom which does sometimes occur in Kant at the expense of the more positive concept. Roughly, the negative reading is freedom from sense-impulse, which can affect free choice but cannot determine it. The positive concept, which I view as more important, is the power of pure reason to be itself practical.³⁶ The metaphysical centrality of this positive concept of freedom is illustrated in this quotation from The Metaphysics of Morals:³⁷

..[I]n reason's practical use the concept of freedom proves its reality through practical principles which, as laws of a causality of pure reason which is independent of all empirical conditions (of sensibility as such), determine choice and prove the existence in us of a pure will in which moral concepts and laws have their source.

This is asserted despite Kant's statement that "Because the concept of freedom is such that no example adequate to it can ever be given in any possible experience, freedom is not an object of our possible theoretical knowledge."³⁸

What may also have encouraged the "dry stick" view is, ironically enough, what strikes me as most personal and revealing in Kant's writing: the struggle against odd inclinations and chance passions, on the grounds of their irregularity, their vulnerability to contingency, rather than their content. Even the virtues are vulnerable--an observation surprisingly similar to Nussbaum's analysis of Hecuba. Moral innocence would be "glorious" but for the high probability of its downfall.

Innocence is indeed a glorious thing; only, on the other hand,

it is very sad that it cannot well maintain itself, and is easily seduced. On this account even wisdom--which otherwise consists more in conduct than in knowledge--yet has need of science, not in order to learn from it, but to secure for its precepts admission and permanence.³⁹

Similarly, accepting favours is a breach of one's duty towards oneself because there is always a risk that the debt will be called in, in some unpalatable and unforeseeable way.

This thrust towards controlling the unpredictable and unstable is a continual theme in Kant, as in this quotation from Religion within the Limits of Reason Alone: "...[W]hen incentives other than the law itself (such as ambition, self-love in general, yes, even a kindly instinct such as sympathy) are necessary to determine the will to conduct conformable with the law, it is merely accidental that these causes coincide with the law, for they could equally well incite its violation."⁴⁰ A similar concern permeates the Grundlegung: "The sublimity and intrinsic dignity of the command in duty are so much the more evident, the less the subjective impulses favour it and the more they oppose it, without being able in the slightest degree to weaken the obligation of the law or to diminish its validity."⁴¹ Because we cannot summon up the emotions at will--not without Pecksniffian hypocrisy, or Hitler's unpleasant power to induce hysteria in himself as well as in his audience--the feelings are not a certain foundation for right action.

Yet very early on in the same work Kant also warns us that hedging one's emotional bets cannot alone constitute the ethical life. "Moderation in the affections and passions, self-control and calm deliberation are not

only good in many respects, but even seem to constitute part of the intrinsic worth of the person; but they are far from deserving to be called good without qualification, although they have been so unconditionally praised by the ancients."⁴² (There is a nice little irony in this last phrase: Kant is disparaging sophrosyne. Effectively this stands Nussbaum's thesis on its head: the ancients now come out as the "dry sticks."⁴³) But the main point here is that Kantian morality can be "beautifully human", that Kant does take account of the full range of human needs and attitudes, does not want us to close ourselves off through an obsessive frugality of the emotions and continual scrutiny of our motives. In particular he is fully aware of our need for a motivating force to obey the Categorical Imperative which cannot be provided by logic alone.

Morality consists in this, that an action should arise from the impulsive ground of its own inner goodness...That it is so is well appreciated by the understanding. Nevertheless, this impulsive ground has no driving force...But we should be on our guard against becoming hypercritical about it, against probing too deeply into its incapacity to attain moral purity. Those who are forever on the look-out for moral impurities in their actions tend to lose confidence in their ability to do good and moral actions...We must rather believe that rectitudo moralis can be a strong impulsive ground of our actions.⁴⁴

I think it is rather shallow to depict Kantian ethics as itself shallow, particularly in its allegedly life-hating narrowness and intolerance of risk. The strictness of the Kantian standard is itself an impressive piece of risk-taking. "Even though there might never yet have been a sincere friend, yet not a whit the less is pure sincerity in friendship required of every man, because, prior to all experience, this duty is involved as duty in the idea of a reason determining the will by a priori principles."⁴⁵ His critics were not the first to recognise the

ensuing risk--that Kantian ethics will produce Tartuffes, as Hegel put it. Kant accepts this danger but still claims that we can know hypocrisy when we see it. The vastness of the risk lies in the overpowering odds that so far as know, we will never encounter anything but hypocrisy: in our inability to be certain, for reasons similar to Karl Popper's arguments about falsifiability, that any actual example is determined by moral motives.

We cannot show with certainty in any example that the will was determined merely by the law, without any other spring of action, although it may appear to be so. For it is always possible that fear of disgrace, perhaps also obscure dread of other dangers, may have a secret influence on the will. Who can prove by experience the non-existence of a cause when all that experience tells us is that we do not perceive it?⁴⁶

This is a Kant to whom Nussbaum should warm: one with some awareness of paradox and pathos, as in his sad assertion that even innocence is bound to decay. Perhaps this is not strictly speaking a tragic awareness, and the usual interpretation, as in White, is that Kant does not believe there are true tragic choices, genuine conflicts of duty. But it has an oddly Greek feel to it, much like Plato's prediction of the inevitable degeneration of the Republic into a lesser political form. "Nothing gold can stay," as Robert Frost puts it. And Kant also recognises that freedom is the genesis of evil as well as of good:

The inherent value of the world, the summum bonum, is freedom in accordance with a will which is not necessitated by action. Freedom is thus the inner value of the world. But on the other hand, freedom unrestrained by rules of its conditional employment is the most terrible of things...All evil in the world springs from freedom.⁴⁷

Indeed, Nussbaum is not so much of an anti-Kantian as she thinks she is. Sometimes she even recognises this openly:

We can claim to be following a part of the deep motivation behind Kant's own view of duty when we insist that duty does not go away because of the world's contingent interventions. Greek polytheism, surprisingly, articulates a certain element of Kantian morality better than any monotheistic creed could: namely, it insists upon the supreme and binding authority, the divinity so to speak, of each ethical obligation, in all circumstances whatever, including those in which the gods themselves collide.⁴⁸

This conflict of duties is the source of tragedy, and it is worth a brief concluding discussion: partly because medical ethics and public policy offer so many examples of tragic choice, such as allocation of scarce medical resources and nuclear deterrence. It may seem that Greek polytheism can have little to do with kidney allocation, and that the discussion of tragedy will be irrelevant. But as Bernard Williams says, "There is no need of irrational gods, to give rise to tragic situations."⁴⁹

Tragedy actually seems to be a source of moral value in Nussbaum, a view which would be rejected by both Kantians and consequentialists. To repeat my opening paragraph, what makes human endeavour admirable to her is its fragility, though what makes my own endeavour valuable is also that it is mine, self-generated, not the product of some external command or goal. A Kantian could agree with the second premise; a consequentialist probably would not; and it does seem to me that there are serious problems in it for an Aristotelian such as Nussbaum. My virtue, as MacIntyre would imply, is so much conditioned by nomos and custom in Aristotle as not to belong uniquely to me. However, I do not intend to press this further.

Nussbaum maintains that to the Greek playwrights at least, tragedy did not necessarily entail paradoxes about moral luck. It is possible for good character to survive tragic outcomes in the sense of external ill-fortune which happens to agents. So Anna Karenina would not necessarily be proved to have made a wrong moral choice by her tragic end. Nor, hard as it sounds, would a case of scarce resource allocation in which the patient given the resource then dies be tragic in the strict sense for the patient who was not given the resource, although it might be tragic for the health professional who makes the decision. (This is because it is the health professional who is the moral agent in this case, not the patient: because if there is an irreconcilable conflict of duties, it is the health care worker who faces it.) Nussbaum argues that what good character cannot endure is having to do things "otherwise repugnant to [agents'] ethical character and commitments, because of circumstances whose origin does not lie with them."⁵⁰ Tragedy is most disturbing when it depicts good people doing bad things. However, they can retain a measure of their goodness and human dignity by recognising that they are doing bad things.

This is something more than a recommendation of hypocrisy. Nussbaum says that what is most evil about Agamemnon's decision to sacrifice Iphigenia is not that he murders her but that he becomes more and more convinced, as the play proceeds, that he should be lauded for killing her. Because the sacrifice brings favourable winds and good (outcome) luck, he is able to ignore the ill-luck (in the situations encountered) which forced him to cause suffering no matter which course he chose.

Agamemnon seems to have assumed first, that if he decided right, the action chosen must be right; and second, that if an action is right, it is appropriate to want it, even to be

enthusiastic about it. From "Which of these is without evils?" he has moved to "May all turn out well."⁵¹

But of course all cannot be well. Agamemnon has killed his daughter. The chorus blames him not so much for her murder as for his callous attitude towards it, his lack of remorse. And without remorse we cannot be said to be either fully human or genuine moral agents. In this sense an awareness that acting well is open to luck but that it is nevertheless incumbent on us to act well--a position with which I think Kant could agree--upholds rather than threatens the ethical enterprise.

Aeschylus has shown us how thoroughly, in fact, the pain and remorse...are bound up with ethical seriousness in other areas of life; with a seriousness about value, a constancy in commitment, and a sympathetic responsiveness that we wish to maintain and develop in others and in ourselves...[W]ithout...acknowledgement of the tragic power of circumstance over human goodness we cannot, in fact, maintain other valued features of our goodness: its internal integrity, its ongoing fidelity to its own laws, its responsiveness of vision.⁵²

This seems to be a sensitive and profound reading by Nussbaum, and I return to it at the end of the dissertation, where I use the Agamemnon example as a parallel for the decision to maintain a strategy of nuclear deterrence. But here Nussbaum upholds a version of ethics with surprising resemblances to Kantianism. Moral personality can remain beyond the reach of external contingency to corrupt, and indeed the power of even evil contingency can be beneficial to an innately good character: pathei mathos. Nussbaum qualifies this assertion in her discussion of Hecuba,⁵³ but she nevertheless maintains that the good moral agent is not one who obtains the good outcome, but one who approaches dilemmas in the right way: compassionately, honestly, imaginatively, with a lack of self-delusion either about the possibility of avoiding tragic outcomes or the likelihood

of their being blessings in disguise. "Aeschylus has indicated to us that the only thing remotely like a solution here is, in fact, to describe and see the conflict clearly and to acknowledge that there is no way out."⁵⁴

But Nussbaum's reading of this supreme example of ill-luck is perfectly consistent with Kant's assertion that what I am forced to do is not my action, since it does not proceed from my choice. Agamemnon was compelled by Zeus to choose between two evil alternatives, and he cannot be condemned within Kantianism for either action. (This is why there cannot be a genuine conflict of duties here, to Kant, since duty is self-generated.) Nussbaum does recognise, though only in a footnote⁵⁵, that Kant could get round the Agamemnon example in this way, though she asserts that "this, however, in no way affects the general point; for he cannot extricate himself from all conflicts in this way." But that is a rather sophistical argument, since Nussbaum has chosen this example to be telling against Kant. Further, Kant may not be able to defuse all conflicts by invoking the notion of compulsion, but the concept does afford one way of limiting the paradox of moral luck. I have not yet asserted that we can eliminate the dilemma altogether, but I do think that we should be looking for such ways of fencing it in.

In any case, perhaps Kant would agree that Agamemnon cannot avoid responsibility altogether in this example. He cannot be held responsible for taking either of the two alternatives which the god has forced on him. But he can be blamed for justifying or even welcoming Iphigenia's death, since he is not compelled to do this.

Another way of delimiting what counts as ethical lies in Kant's inveighing against "fantastic virtue", his assertion that some matters are matters of moral indifference.⁵⁶ There is no moral merit in ostentatiously multiplying the occasions of duty. In this sense, Hegel is wrong to accuse Kant of giving rise to Tartuffes.

Some burden their consciences with many matters of negligible importance...Conscience should not lord it over us like a tyrant; we do no hurt to our conscience by proceeding on our way cheerfully; tormenting consciences in the long run become dulled and ultimately cease to function.⁵⁷

Kant might have argued that the conditions which make Williams's sort of scepticism about ethics possible also make it false: self-consciousness, the awareness of one's actions as something other than conditioned. Williams says that the doubts arising from the perception that ethics contradicts luck are ultimately at risk of shrinking the idea of agency to an extensionless point; yet without the idea of agency there could be no questions about moral luck. The condition which makes the moral luck paradox possible also disproves it: the existence of agency. Yet the unconditioned freedom of the moral agent remains paradoxical. In this sense it is both right and wrong to assert that Kant tries to erase the paradox of moral luck. But it is certainly ill-advised to depict him as wilfully blind to it.

Practical reason is the cornerstone of our selfhood in Kant, what makes us "beautifully human". Although it is not the source of character, on which constitutive luck and events take their toll, it is the source of personality in Kant's sense of the term. This is best expressed in the

moving conclusion to the Critique of Practical Reason:

Two things fill the mind with ever new and increasing admiration and awe, the oftener and more steadily we reflect on them: the starry heavens above me and the moral law within me. I do not merely conjecture them and seek them as though obscured in darkness or in the transcendent region beyond my horizon: I see them before me, and I associate them directly with the consciousness of my own existence...The former view of a countless multitude of worlds annihilates, as it were, my importance as an animal creature, which must give back to the planet (a mere speck in the universe) the matter from which it came, the matter which is for a little time provided with vital force, we know not how. The latter, on the contrary, infinitely raises my worth as that of an intelligence by my personality, in which the moral law reveals a life independent of all animality and even of the whole world of sense--at least so far as it may be inferred from the purposive destination assigned to my existence by this law, a destination which is not restricted to the conditions and limits of this life but reaches into the infinite.⁵⁸

The surprising similarity between Nussbaum's and Kant's arguments can be seen even more clearly in contrasting Nussbaum's view with an outcome-orientated one such as utilitarianism. After all, what is evil in Creon, to Nussbaum, is that he considers "only" the welfare of Thebes, defining right and wrong purely in outcome terms. It is to utilitarianism and moral luck that I will now turn.

FOOTNOTES

1. Nussbaum, Fragility of Goodness, p. 3. In a critical review (see footnote 13), Nicholas P. White queries whether this link between the vulnerability and the value of human endeavour may be of more concern to Bernard Williams, whose influence Nussbaum acknowledges, than it ever was to the Greeks.

2. Nussbaum, p. 3. As I did in the first chapter, Nussbaum refuses to equate luck with randomness; by it she simply means external contingency, everything which occurs through the agency of someone or something other than the moral agent.

3. Ibid., p. 2, original emphasis.

4. This seems a startling proposition, even though Nussbaum is careful to distinguish the influence of Kant in his pure and subtle form from that of vaguely styled Kantians. Nevertheless, it says little about the far greater influence of consequentialism in Anglo-American philosophy, at least until the last two decades, perhaps.

5. Ibid., p. 4.

6. Alasdair MacIntyre, After Virtue (London: Duckworth, 1981).

7. Nussbaum, Fragility of Goodness, p. 5.

8. Immanuel Kant, Lectures on Ethics (tr. Lewis White Beck) (Harper Torchbooks, 1963), "Conscience", p. 132.

9. Thomas Nagel, The View from Nowhere (Oxford: Oxford University Press, 1986), p. 195 ff. In addition to identifying five possible combinations of the moral and the good lives where Nussbaum and Williams locate only two, Nagel introduces a third variable which could furnish further permutations: the rational. However, I shall leave off discussing that until chapter four. Williams's dichotomy is similar to that described in my chapter one, in relation to his Ethics and Limits.

10. Ibid., p. 194.

11. Ibid., p. 196.

12. John Rawls, A Theory of Justice (Cambridge, Massachusetts: Harvard University Press, 1971). Rawls of course describes his work as "highly

Kantian in nature" (p. viii).

13. White, "Rational Self-Sufficiency and Greek Ethics," p. 141.

14. Ibid., p. 140.

15. Cooper, review of Nussbaum in Philosophical Review, p. 560.

16. Beck, preface to Kant, Lectures, p. xii.

17. White, "Rational Self-Sufficiency and Greek Ethics," p. 143.

18. Nagel, "Moral Luck", p. 24.

19. Nussbaum, Fragility of Goodness, p. 399.

20. Ibid., p. 63.

21. See, for example, Laurence Blum, Friendship, Altruism and Morality (London: Routledge Kegan Paul, 1980). Blum produces a more sophisticated version of this criticism of Kant as a dry old stick than is usually found.

In his view, Kant's concentration on the fortitude of the good will against any possible adversity leads him to ignore the greater general reliability of the good emotions, despite their vulnerability to external contingency.

In the statistical long run, they are more reliable in producing the correct action than is Kantian beneficence. "If what we want from a motive is that it in fact leads us to morally good acts, we should be more interested in the extent to which the given motive actually does this, in the face of the counter-indications which we as human beings actually face, than in whether the motive is capable of withstanding any possible counter-motive." (p. 31)

22. Nussbaum, Fragility of Goodness, footnote 7 to page 52, discussing Antigone.

23. Roger Scruton, Kant (Oxford Past Masters Series, 1982), pp. 72-73.

24. Kant, The Metaphysics of Morals, part two (tr. Mary J. Gregor as The

24. Doctrine of Virtue (Philadelphia: University of Pennsylvania Press, 1964), 221.
25. Kant, Critique of Practical Reason (tr. Lewis White Beck) (Indianapolis: Bobbs-Merrill, 1956), preface, 4, footnote 1.
26. Scruton, Kant, p. 64: "Freedom is the power to will an end of action for myself...If my action is called unfree it is because there is a sense in which it is not truly mine."
27. Lectures, "Duties to Oneself", p. 118.
28. MacIntyre, After Virtue, p. 119.
29. Lectures, "Responsibilities for Consequences of Actions", p. 59..
30. Ibid., p. 60.
31. Scruton, p. 66.
32. Lectures, p. 60.
33. Ibid., p. 62.
34. Nagel, "Moral Luck", p. 29.
35. Critique of Practical Reason, 8, footnote 5.
36. Metaphysics of Morals, part two, 213.
37. Ibid., 220.
38. Ibid., 220.
39. Kant, Fundamental Principles of the Metaphysic of Morals (Grundlegung zur Metaphysik der Sitten) (tr. Thomas K. Abbott) (Indianapolis: Bobbs-Merrill, 1949), 26, emphasis added.
40. Quoted in Blum, p. 26.
41. Grundlegung, 52.
42. Ibid., 12.
43. Nussbaum discusses sophrosyne in connection with the Phaedrus (pp. 204-

6) but not in gender-specific terms. Indeed, the object of Lysias's exhortation in its favour is male. It is interesting that Nussbaum views the dialogue as urging the young man ultimately to venture outside the safe walls of sophrosyne (p. 223), just as the setting of the action is outside the city walls--uniquely in the Socratic dialogues.

44. Lectures, p. 65.

45. Grundlegung, 30.

46. Ibid., 45.

47. Lectures, p. 122.

48. Nussbaum, Fragility of Goodness, p. 49, original emphasis.

49. Williams, "Ethical Consistency," in Problems of the Self (Cambridge: Cambridge University Press, 1973), 166-186, at p. 173.

50. Nussbaum, p. 25.

51. Ibid., p. 36.

52. Ibid., p. 50.

53. An interesting further application of Nussbaum's discussion of Hecuba can be found in Card, "Women's Voices and Ethical Ideals". In criticising the influential work of Carol Gilligan (In a Different Voice, Harvard University Press, 1982), who argues that men and women develop separate but equally valid ethical systems, Card uses Hecuba as an example of the dark side of separate development, the vices endemic to the powerless. This is presented as a form of moral luck.

54. Nussbaum, Fragility of Goodness, p. 49.

55. Ibid., chapter two, footnote 64, p. 434.

56. Metaphysics of Morals, part two, 409. However, at 222 Kant seems to leave this question open.

- 57. Lectures, pp. 134-135.
- 58. Critique of Practical Reason, 161-162.

Chapter Three.

Probability and Utilitarianism

Although I have not presented Kant as striving to create a risk-free system of ethics, I do not deny that moral luck looks at first to be particularly paradoxical for a Kantian. A Kantian is primarily concerned about the state of his or her own will; but how can a will be stably good if its goodness is subject to factors beyond the agent's control? It might be easier for the good will to "sparkle like a jewel in its own right, as something that had its full worth in itself," if it were the only thing in the world, "if there remained only the good will." It is the subjection of the will to the outside world which creates the dilemma of moral luck. Nussbaum can be seen as arguing that Kant deliberately chose to ignore the way in which the agent simply is subject to chance in order to preserve the integrity of the good will. I have asserted that this is a misreading of Kant, and that Nussbaum's own position is closer to Kant's than she acknowledges. And I have argued that Kant both recognises the moral luck paradox and suggests a resolution of it, through limiting what agents are responsible for whilst retaining a strict notion of agency. Agents are not responsible for things they are compelled to do; nor should the bounds of the ethical be stretched to include "fantastic virtue", to make too many matters moral.

Because utilitarians emphasise consequences rather than conscience, they seem to have no such difficulty to get over at all, however, and the

dilemma of moral luck appears not to exist for consequentialists in general. It is only by asking Kantian sorts of questions that it arises. Perhaps this might be a good time to restate the paradox, which depends on a dichotomy between the outside world and what is defined as the locus of moral worth in the individual agent. If good and bad pertain purely to consequences, rather than to the will, it looks at first as if the problem of moral luck cannot arise. The dichotomy should no longer exist because the locus of moral value has shifted from the agent to the outside world. Since the paradox of moral luck seems to result from this tension between individual responsibility and determination by things outside the agent, consequentialism should be able to eliminate the paradox.

Common-sense judgment holds people responsible for their actions according to how those actions turn out; yet common sense also maintains that people are not responsible for matters beyond their control. Much of how things turn out is inevitably beyond agents' control. It is possible to resolve this, the paradox of moral luck, by ceasing to hold people responsible for what is beyond their control, but then we may not wind up holding them responsible for very much. This is essentially the line which Williams has pursued, with increasing severity over the years--so that a fixed ethical system becomes more and more of a nonsense in his recent work, to be replaced by a consensual system of societal values. It is not the path which I intend to take in the rest of this thesis.

Instead I mainly want to investigate the alternative way of resolving the paradox, closer to Nagel's formulation. This is to retain a

broadly Kantian idea of agent responsibility despite the actions of chance, but to limit what agents are responsible for. Recall, for example, that Nagel did not think that aesthetic failure would have made Gauguin's decision morally wrong; nor would success as an artist have proved his choice ethically correct. Even if success "still determines his most basic feelings, that shows only that his most basic feelings not need be moral."¹ Chapters five through eight of this dissertation comprise this effort to limit what agents are responsible for, so as to avoid the moral luck problem, by illustrating how such limits might apply in some areas of practical ethics. They may appear at first to focus not so much on moral luck as on how to make the optimal choice under conditions of uncertainty, but I shall deal with that objection when I come to chapter five.

Meanwhile, let us return to consequentialism and moral luck. This chapter has quite a modest aim: to argue that moral luck is not a peculiarly Kantian fixation, even though the dilemma appears to arise by asking Kantian questions. Again, it seems at first that consequentialism does not have a problem with moral luck, though of course it does concern itself with how to make the optimal choice under conditions of uncertainty. I want to argue, however, that what looks like a split here is not genuine, and that moral luck is relevant to consequentialism. The reason why this is true relates to how consequentialism has defined what kinds of outcomes the agent is responsible for, and this in turn relates to probability. The rest of this chapter will concern probabilities and outcomes in consequentialism, particularly in utilitarianism.

To ascribe responsibility on the basis of consequences inevitably raises the question of whether the agent is responsible for the actual or the potential consequences of his or her action. If it is the actual consequences, we seem not to have evaded the problem of moral luck after all. The rightness or wrongness of my decision will still depend on some matters beyond my control, if matters do not turn out as I had intended or foreseen. (I shall not distinguish the two at this point in the argument.) If we still want to retain some conception of responsibility and agency, this creates a temptation to think in terms of potential consequences, which is normally translated into the likelihood of those outcomes and their desirability, or probability and utility.

Whether it is actual or probable consequences which are most ethically relevant is a debate which goes back at least to Russell and Moore, and which is summarised by C.D. Broad in "The Doctrine of Consequences in Ethics."² To be more precise, Broad thinks that the Moore-Russell controversy suggests that even ascribing responsibility by actual consequences requires probabilities to be taken into account. Although Moore held, against Russell, that it was actual consequences which mattered, he was compelled to introduce probability in the end, Broad asserts. But this is not the main point at the moment, however: rather, the question is whether a consequentialism that thinks solely in terms of actual consequences--whether or not it has to incorporate probabilistic reasoning--is not just as vulnerable to the paradox of moral luck as Kantianism looked to be.

It does appear that Russell was aware that reliance on actual consequences as the touchstone somehow lessens the agent's control and may make it hard for us to know just what the agent could be responsible for. This is essentially the moral luck problem. Using probable rather than actual consequences as the guide has the advantage, as Broad phrases Russell's argument, of "making objective rightness independent of unforeseeable circumstances."³ Now not controlling is broader than not foreseeing: foresight is at best a minimal sort of control. I will need to distinguish the two later on in this chapter, and again in the next chapter, on rationality. But there is a germ of a moral luck argument here, and one which appears to work in favour of potential rather than actual consequences.

In contrast, Moore is not too bothered by the dilemma of moral luck. Being judged by the actual consequences, he asserts, will not turn out to be intolerably harsh, because we ought to distinguish what is right to do from what is right to praise. Further, rightness only refers to consequences, according to Moore, but we may still say that the agent's choice was good even if the consequences turned out "wrong". Although Broad calls Moore's formulation generally more plausible than Russell's, his preference for Moore's position does not extend to his claim about praise and blame:

This supposition is not necessarily true. A's praise or blame of B's act is a second act, and like all others, its rightness or wrongness must be judged by its own consequences, and not by those of B's act.⁴

So Broad is unwilling to allow Moore this evasive action. But this may be because Broad himself does not fully recognise how deeply the moral luck

problem would affect a consequentialism based on actual outcomes. He is blithely willing to accept that "since the rightness of your action is at the mercy of all that is going to happen in the universe throughout all future time, there is no reason to expect better results from conscientious acts than from the most stupid and biased ones."⁵ This seems hopelessly demoralising: to a consequentialist, the point of making conscientious decisions is precisely to obtain desirable outcomes. If there is no reason to expect better results from conscientious acts than from stupid ones, there can be no reason in consequentialism to make one's ethical choices in a conscientious manner. (There would be reasons in Kantianism, of course.) The first half of Broad's claim, at least, might strike Nussbaum as an admirably tragic-minded view, but it is hard to see how Broad's formulation could motivate right action.

Broad's main reservations about relying on potential outcomes concern Russell's alleged mingling of actual and probable consequences, with disastrously confused results:

It is not clear whether the objective rightness of an act depends on the actual value of its probable consequences, or the probable value of its actual consequences, or the probable value of its probable consequences. All we are told is that it does not depend on the actual value of its actual consequences.⁶

But of course this does not defend the actual-consequences position against the moral luck problem; it merely criticises Russell's particular formulation of the potential-consequences line. Is it impossible to formulate a consequentialism of actual outcomes without encountering moral luck dilemmas? I have suggested so far that the answer is probably yes, but A.N. Prior's work provides additional grounds for doubting that an

actual-consequences utilitarianism can work.⁷ His argument bears on moral luck, though it is not couched in moral luck terms: rather, in those of a different paradox.

Moore maintains that it is our duty to perform that action which actually produces the best total consequences, of all available alternatives. But Prior says that it is a practical impossibility to determine the contents of that duty. This is true not only because we lack the power to predict; it is also because Moore is making paradoxical assumptions about control. No such best package of outcomes exists unless determinism is total; but of course if determinism is total and free will non-existent, we cannot be said to make ethical choices. The total future of the world depends on how other agents choose, too: not just on what I choose.

Now it might be possible to approach such decisions through game theory, with the game being viewed as played against others rather than against a neutral nature. But whether they could be approached through ethical reasoning, particularly consequentialist reasoning, is more problematic.⁸ One consequentialist attempt to get away from the notion of nature (stepmotherly or otherwise) as the sole determiner of outcomes has been made by Bart Gruzalski, although in a manner rather incidental to his main concern, "the defeat of utilitarian generalization."⁹ Gruzalski does indeed tend to present future outcomes as knowable: perhaps controllable, and certainly very precisely knowable, quantifiable in probability terms--when in fact they depend on the decisions of all agents. But these

decisions are meant to be influenced by the very arguments being made on the basis of this future package of outcomes. This entails a profound circularity.

Gruzalski has devised the hypothetical case of a worker, Fred, whose decision about whether or not to strike is dependent for its outcome's success on other workers' decisions. (Gruzalski does not debate whether this is an ethical or merely a prudential decision, but again, I leave such questions until the following chapter.) The conditions surrounding Fred's choice are as follows:

A strike has been called. If the strike fails, the consequences would be very undesirable, whereas the consequences of a successful strike would be very desirable. Everyone knows that almost certainly all 450 workers will strike and that the strike would still be successful if only 300 workers were to strike. Although each act of striking produces an undesirable consequence (loss of salary), these consequences, even if magnified five times, are less undesirable than the strike's not succeeding. In these circumstances, what should a utilitarian do?¹⁰

What makes this calculation problematic is that Fred's action changes the probabilities of success. In fact, this is the classic free-rider problem.

A recent economic study of the free-rider dilemma has concluded, interestingly, that it cannot be solved except through a Kantian system emphasising beneficence rather than self-interest.¹¹ The best outcome for Fred would be to keep his pay packet and to see the strike succeed. But if everyone reckons like Fred, the strike will fail. However, even if Fred reports for work, the strike will still succeed provided that no more than 149 other workers reason as Fred has.

Yet no one--including Gruzalski--can know in advance how the other workers will reason. Gruzalski simply stipulates that the strike's probability of success is very high, 0.9999. Now this is not actually a mistake in moral mathematics, although Gruzalski does make such errors: for example, in one table a probability of 3,550 appears, when probabilities can only range between 0 and 1. Gruzalski can of course stipulate whatever probability of others' striking he likes; but if we are to regard Fred's decision as generalisable and instructive, we have to view all the workers as Freds, and all their decisions as up for grabs. It is therefore puzzling when Gruzalski says that "everyone knows that almost certainly all 450 workers will strike." After all, Fred is one of the 450, but his decision is regarded as undetermined. And "everyone could know" equally well--indeed, perhaps more certainly--that all the employees will turn up for work because each wants to be a free-rider.

Matters become even more tangled when Gruzalski tries to derive Fred's likely share in the strike's success, if it succeeds, from the probability that 399 other workers will actually strike in the end. This appears to contradict his earlier assignments of probability.¹² If it took 300 workers to assure the strike's success, we only know that at least 300 stayed home; how did the 400 creep in? Once again Gruzalski has simply inserted it as an a priori proposition. Again, in a sense this is perfectly legitimate: Gruzalski can do what he likes with his own example. He defines his, and any utilitarian's concern as being "with what a person or group of persons should do given the facts, not what should be done were the facts different."¹³ But in real life the question is often not what to do with

either these or different facts, but what to do with no facts. How can Fred know in advance what his likely share in the strike's success will be? Can he control his co-workers' decisions in any way? Certainly he needs at least to know his likely share of credit in order to decide what to do, in order to balance the gain attributable to his action against the loss of pay.

Again, there is a distinction between the uncertainty introduced by our inability to control many of the decisions of others and that introduced by our lack of knowledge of the future. The first is certainly part of the paradox of moral luck, but the second also enters in. It is the first that Prior regards as the problem for a consequentialism of actual outcomes. The second, Prior says, "is only part of the general problem of 'duty and ignorance of fact', which has nothing specially to do with utilitarianism, and [which] was allowed for by Moore anyway."¹⁴ However, it is debatable whether Fred's ignorance of fact can actually be shrugged off this lightly. In any case, lack of control over outcomes is certainly difficult to get round, though in a reply to Prior, D.D. Raphael attempts to argue that "the consequences of my actions are also the consequences of yours, [in] that we both bring (in the ascriptive sense) whatever happens."¹⁵ If I am seen as having some control over the consequences of my actions, I also have some over yours. However, even Raphael admits that this is "stretching the meaning of the concept of action and its consequences far beyond conventional use."

How does this bear on moral luck, and what happens now to the actual-consequences variety of utilitarianism? Recall that the dilemma of moral luck concerns the simultaneous common-sense requirements that we be held responsible only for actions which we can control and that we be held responsible for our actions according to how they turn out. If Prior is right, a doctrine of actual consequences radically extends what we are responsible for to include the decisions of others; but these we cannot normally control anything like reliably. This might result in what Williams calls "the final destruction" (in a slightly different context):¹⁶ the combination of Kantian strictness of duty with an impossibly broad and (according to Prior) ultimately undefinable list of duties to fulfill. I shall argue in the rest of this dissertation that the best way to avoid this worst of both worlds is to retain a fairly strict notion of agency but to limit equally strictly what agents are responsible for.

Prior prefers to leave agents responsible for a much broader range of outcomes, but not for actual consequences. We may have to be content to consider as our duty that action which will probably engender the most desirable total consequences, he says. But this too leads to problems: for example, probabilities must be conceived as having an objective existence, Prior stipulates. It is not enough that we believe the consequences of an action are likely to a certain degree; they must actually be probable to that level of magnitude. After all, what affects other people is the outcome, not my estimate of it.¹⁷ However, one obvious difficulty is exactly what probabilities represent. As Broad says about the related concept of expected value (probability times utility), "[there is] no

reason to think that the notion of mathematical expectation is really a measure of anything in this world."¹⁸ This is true mainly because its component of probability may likewise have no substance: utility represents a definite something desirable, but probability may not represent anything at all. This is not the least of the problems and paradoxes associated with probability, and it is to these, and the associated probable-consequence variety of utilitarianism, that I shall now turn.

I said before that we need to distinguish uncertainty resulting from ignorance with that stemming inevitably from the freedom of other agents. A concentration on probable consequences looks at first as if it can lessen much of the first kind of uncertainty, through the statistical discipline of probability interpretation. (However, this does not answer Prior's objections about control of others' actions.) And of course we do incontestably use probabilities in ordinary prudential decisions. Why not also use them in moral choices? I do not intend to tackle any such broad question in this chapter; in the next chapter I do give further consideration to the distinction between prudential and ethical reasoning. But aside from the arguments within statistical theory about whether probabilities are objective or subjective--and as I have indicated in discussing Broad's doubts about whether mathematical expectation represents anything real, such questions about the status of probabilities are substantial matters, not just technical questions--there are grave questions about using probabilities as the touchstone for right action.

Two sorts of questions arise. The first is to do with mistakes in moral mathematics, the second with incommensurabilities of moral value. Both sorts are illustrated in the following example, but I shall need to separate them out in my elucidation of it. Broadly, however, the point of the example is that we cannot improve matters by trading in one paradox for another. I want to suggest that the paradoxes of probability are no less crippling than those which Williams believes that moral luck presents, and that some utilitarians have been insufficiently aware of their imperviousness to better methods of calculation.

Consider a typical probability, that of dying in a motor vehicle accident. The risk for Americans in 1975 was .0027, or 27 deaths for every ten thousand Americans per year.¹⁹ These deaths constitute a statistical certainty, within the limits of the appropriate degree of confidence. Of every ten thousand Americans, 27 could "certainly" have expected to die in road accidents that year, although of course no one knows which 27 would die. But the group total is certain, and indeed is derived by taking those who did die in road accidents that year as a proportion of the populace. Now consider a probability derived with similar certainty from actual figures, though in a hypothetical example.²⁰ The point here is that there is no quibble about whether the probability is accurate, and no improvement in the paradoxes presented by the example could be made by obtaining the initial probabilities in some more accurate fashion. There is no more accurate fashion. Nor are there the game-theoretical sorts of problems I identified in Gruzalski's example.

Twenty-five prisoners are exercising in a yard, under the surveillance of a lone guard. A solitary witness, who is too far away to identify any of the prisoners, sees the guard--recognisable by his uniform--trip and fall, knocking himself out. After huddling together for a moment, the prisoners separate. One hides in a shed in the corner, whilst the other twenty-four fall on the guard and kill him. Then the twenty-fifth man re-emerges and mixes with the other prisoners. The rest of the prison guards rush into the yard, where they find the dead man and the twenty-five prisoners. How many, if any, will be found guilty of murder?

None, says the originator of the example, Charles Nesson: no jury would convict any of the prisoners. To do so would be equivalent to announcing that it believed a probability of .96 (24 in 25) to constitute guilt "beyond a reasonable doubt," the standard in criminal prosecutions. This would be detrimental because doubt is functional: the public can only defer to jury verdicts if they are not couched in such bald and cynical terms. This is not actually the point about value which I want to discuss; I shall not pursue this functionalist line. Instead, I want to ask whether the problems here arise from the moral imperative that we should not convict someone who might be innocent on a probabilistic basis, or from mathematical difficulties. Difficulties there are in plenty: contradictions occur whether the jury convicts all twenty-five prisoners, twenty-four, or none.

If the jury convicts the first prisoner in the dock on the .96 probability of his guilt, it must do so with all the remaining prisoners.

But that is equivalent to announcing that twenty-five men killed the guard, which we know to be false. That is, at the end of all the trials the probability of each man's guilt would have been proclaimed to be 1.00. But if the jury convicts no prisoners, it announces that the probability was 0.00 for each man. The only way the jury could obtain the statistically correct result for the group as a whole would be to convict any twenty-four men on any basis it liked, randomly or systematically. But somehow this is even more abhorrent than convicting all twenty-five.

Here is where the distinction between ethical paradox and mathematical error comes in. It looks as if we can solve the mathematical problem by convicting twenty-four prisoners, chosen at random. (In a moment I shall note that this would only produce the statistically correct result for the group, however.) But no one would assert that this will solve the ethical problem. The jury will have transgressed the basic Anglo-American legal principle of judging each case on its merits, without a predetermination of guilt. Whatever it decides about case twenty-five will have been determined already by whether it chooses to locate the one innocent man among the first twenty-four. And of course there is no guarantee that the man chosen for acquittal will be innocent: indeed, there is a .96 probability against it. Yet the jury will effectively have declared that his likelihood of guilt is 0.00, and that of the others 1.00. In this sense the ethically wrong thing to do will not even have solved the mathematical problem. It will have produced the statistically correct result for the group, but not for any individual.

Whether the jury would be to blame if it chose the twenty-four guilty men arbitrarily has nothing to do with the level of potential loss in the case. It would still have acted wrongly whether or not the death penalty were in force, for example. This is a case about probabilities and their paradoxes, not utilities. To convict, the jury should have to be able to assert that it was incapable of being surprised by a wrong result. Yet it also knows that in each case it stands a .04 chance of being wrong, and that for the group of twenty-five men it stands only a .04 chance of getting the distribution of guilt and innocence right. It has been said that probability would be an appropriate tool "for any legal system that aimed solely to discover the truth more often than not,"²¹ but not for one which aims to dispense justice to individuals. However, a .04 chance of getting the distribution right for the group hardly makes it likely that even that minimal criterion will be satisfied.

If we want to maintain a strict legal principle such as "innocent until proven guilty," we will simply have to ignore probabilities altogether in this case. This will limit our responsibility as agents: we will be responsible for not infringing the presumption of innocence, but we will not be responsible for getting the outcome right in terms of identifying the one innocent man and convicting the others. In fact, we will have to accept that the rightness or wrongness of the jury's decision cannot depend on how things turn out, on outcome luck. A consequentialist ought to be much more bothered about this example than a Kantian, who would be more likely to adhere to a strict a priori principle for its own apparent sake, or rather for the sake of respect for persons. And the

example casts at least some initial doubts on probability as a guide for ethical action.

The case illustrates two points. First, even indubitably accurate probabilities entail statistical problems of the sort highlighted before, the extreme unlikeliness of getting it right for the group and getting it right for the individual. Second, getting it right for the group has to be balanced against another, indubitably ethical value, that of not wishing to convict individuals solely on a probabilistic basis if that means trespassing on the presumption of innocence. Possibly a very extreme consequentialist would reply that there is nothing particularly special about the presumption of innocence, any more than any other a priori principle. This would be consistent, in the sense that the principle "innocent until proven guilty" comes out of someone else's toolkit, and is not an instrument which consequentialists are obliged to use. But it would be a very extreme reading which would step outside the established parameters of the case and the actual ethos of the courtroom. I think that the balancing can only go one way, against using probability as a guide to judgment in this case.

Now some utilitarians have talked as if having a better theory of probability or a better set of techniques for handling it could eliminate all paradox and all incommensurability of values, making utilitarianism unassailable. J.J.C. Smart has written: "What utilitarianism badly needs, in order to make its theoretical foundations secure, is some method according to which numerical probabilities, even approximate ones, could in

theory, though not always in practice, be assigned to any imagined future event."²² The argument seems to be that if we could be sure that we were dealing with objective frequencies rather than subjective probabilities, we could use probabilistic estimates with confidence as the guide to ethical choices. The problem is conceived as admitting of a "technical fix". Smart appears to believe that subjective probabilities only need a little tidying up to make them into the objective probabilities on which utilitarianism can find a firm foundation:

We need a method of assigning numbers to objective, not subjective, probabilities. Perhaps one method might be to...define objective probabilities as the subjective probabilities of an unbiased and far-sighted man. I do not know how to do this, but I suspect, from the work that is at present being done on decision-making, that the situation may not be hopeless. But until we have an adequate theory of objective probability, utilitarianism is not on a secure theoretical basis.²³

However, we have already seen that indubitably objective frequencies cannot always eliminate difficulties in moral mathematics or resolve incommensurability of values, in the example of the twenty-five prisoners. So confidence in a technical fix is misplaced.

Furthermore, subjective probabilities are not merely a downmarket version of objective ones, as Smart implies. To a subjectivist in probability theory, probabilities can never constitute anything but degrees of belief. Therefore probabilities are "assigned" rather than "discovered." A subjectivist would actually be willing to accept a rather Kantian premise: that what I am responsible for is not the outcome of my decision, but the diligence with which I make it. This again constitutes a limiting of what we are responsible for, interestingly enough, but this is really only an aside.

The more important point about subjective probabilities is that conceiving of probabilities in this manner lands us back with the problems which Broad identified: what sort of reality do such estimates possess? None, according to one subjectivist: "The calculus of probability can say absolutely nothing about reality."²⁴ Now it is hard to see how subjective probabilities, construed in this manner, can help a utilitarianism which concentrates on probable outcomes. After all, utilitarianism traditionally claims to be hard-headed, "realistic", and positivistic; this is exactly its defence against luck and chance, as Bentham made clear. "Every circumstance by which the condition of an individual can be influenced, being remarked and inventoried, nothing...[is] left to chance, caprice, or unguided discretion, everything being surveyed and set down in dimension, number, weight and measure."²⁵

I do not mean to argue that we should not try to avoid mistakes in moral mathematics: only that doing so will not always eliminate the paradoxes of probability or resolve problems of incommensurable values. The Nesson example illustrates both sorts of problem, even though it contains indubitably objective probabilities and entails no game-theoretical difficulties about controlling or predicting the actions of other agents. It is not sufficient simply to make sure that we are dealing with objective probabilities if we want to make probabilistic estimates of potential consequences the touchstone of right action. However, it would certainly be crucial for a consequentialism of probable outcomes to avoid whatever mistakes in moral mathematics can be avoided. Before concluding I shall briefly illustrate a class of case in which I do think that it might shore

up this type of consequentialist argument to deal with a factor which is often overlooked when dealing with probable outcomes, statistical independence. This appears to be a case about utility rather than probability, or perhaps about whether rights are additive. But a probable-consequences utilitarianism would also need to use statistical independence in order to handle it successfully.

The sort of case I have in mind is the "innumeracy of ethics" variety, and the example I shall use is the one which first sparked the debate, Elizabeth Anscombe's lifeboat hypothetical.²⁶ (I have altered it slightly to parallel a case developed by Derek Parfit.) Five people are stranded on one rock, and one person on another. There is only enough time for the rescuer to go to one rock: which should he choose?

Anscombe concludes that the rescuer is not wrong to save the single person: the others would have no grounds for complaint so long as the choice is not made for some bad reason, "some ignoble contempt." The rescuer "acts badly if he chooses to rescue rich people rather than poor ones, having ill regard for the poor ones because they are poor. But he doesn't act badly if he uses his resources to save X, or X, Y, and Z, for no bad reason, and is not affected by the consideration that he could save a larger number of people."²⁷ This looks like a deontological argument about a pure will, as Anscombe phrases it. But the decision to save one person rather than five has also been defended in consequentialist terms by John Taurek. It is not necessarily preferable to save many lives rather than one, since, Taurek asserts, we cannot understand how pain could be

additive. This has been vigorously rebutted by Derek Parfit: the deaths of the five would be a worse outcome than a single death. What is interesting about Parfit's rebuttal is that he uses Anscombe's warning against "ignoble" reasons in order to come to the opposite conclusion from Anscombe:

If we give the rich priority, we do not give equal concern to saving each. Why do we save the larger number? Because we do give equal weight to each. Each counts for one. That is why more count for more.²⁸

"Each counts for one." This seems incontrovertible: it should also mean that the five deaths should not be considered as a whole, but as a run of five independent trials. Let us change the probability of effecting a successful rescue from 1.00 to any lesser figure, even 0.9999. That is, let us incorporate uncertainty, either in the situation itself or in the rescuer's mind. Surely this makes the example more realistic? If any uncertainty at all is allowed in the lifeboat example, statistical independence and the calculus of probability now dictate going to the more crowded rock. Let us assume that the weather conditions, the distance of the two rocks, and the limits of the rescuer's strength are such that the chances of reaching either rock are one in five, a probability of 0.20 of successful rescue. The two possible rescues between which the choice is to be made are statistically independent events. Following an expected-value strategy, the rescuer stands to save $1/5$ of a life (0.20 probability times a utility of one) if he goes to the rock with the single person on it, but one life (0.20 times a utility of five) if he chooses the rock with five people. So if numbers come into it at all, the rescuer should try to save the greater number, as Parfit asserts; but to make the numerical reasoning

correct, the argument needs to incorporate statistical independence.

Recognising outcome luck requires that.

Let me recapitulate the argument of this chapter. It seemed at first that consequentialism could avoid the dilemma of moral luck because it emphasises outcomes rather than the good will. But consequentialists have been unable to agree on whether actual or potential outcomes are the standard by which to judge actions.²⁹ There are serious reasons to doubt whether reliance on actual consequences avoids the paradox, since actual consequences are often outside agents' control. This difficulty is cited by Russell, along with the related moral-luck question of what agents can then be considered responsible for, as an argument for emphasis on potential consequences instead.

Using potential or probable consequences as the touchstone appeared to give us greater foreknowledge, if not necessarily greater control over the actions of other agents. But while probabilities can be a source of some enlightenment if we avoid mistakes in moral mathematics (such as ignoring statistical independence, or failing to delineate risk for individuals from statistical certainty for groups³⁰) they cannot do everything which has been claimed for them by some utilitarians, such as Smart. In particular, no amount of improvement in assigning or manipulating probabilities can get rid of conflicts of incommensurable values, such as occurred in the twenty-five prisoners example. Of course we would like to get the outcome right, apportioning acquittal and punishment correctly; but we must not convict an innocent man. In this case it seemed that the only solution was

not to hold the jurors responsible according to whether they reach the factually correct verdict, identifying the one innocent man and convicting the twenty-four guilty ones. Only by not making jurors responsible for getting the "correct" outcome could we avoid being intolerably surprised by how the outcome turned out. Only by limiting what they were responsible for--to upholding the presumption of innocence--was it possible to avoid the depredations of moral luck in this case.

Limiting what agents are responsible for eliminates the problem of moral luck in the twenty-five prisoners example. As I said earlier, narrowing down the ethical requirements in this way helps to avoid what Williams calls the ultimate destruction, the conjoining of a strict Kantian sense of duty with a broad definition of where our duties lie. In the next chapter I want to explore this sort of limiting in more detail. In particular, I want to look at where the ethical leaves off and the prudential begins.

FOOTNOTES

1. Nagel, "Moral Luck", p. 28, footnote 3.
2. G.E. Moore, Ethics (London: Oxford University Press, 1912); Bertrand Russell, Philosophical Essays, rev. ed. (London: Allen and Unwin, 1966);

and C.D. Broad, "The Doctrine of Consequences in Ethics" (1914), reprinted in D.R. Cheney (ed.), Broad's Critical Essays in Moral Philosophy (London: George Allen and Unwin, 1971), pp. 17-42..

3. Broad, "Doctrine of Consequences", p. 27.

4. Ibid., p. 29.

5. Ibid., p. 35.

6. Ibid., p. 26.

7. A.N. Prior, "The Consequences of Actions," in Proceedings of the Aristotelian Society Supp. v. XXX, 1956.

8. Advocates of a game-theoretical approach as a tool for ethicists generally distinguish ethics from game theory, and prudence from ethics. For example, R.B. Braithwaite (The Theory of Games as a Tool for the Moral Philoospher, Cambridge, 1963) separates first-order moral principles from second-order prudential ones and reserves game theory for elucidating the latter.

9. "The Defeat of Utilitarian Generalization", Ethics, v. 93 (October 1982), pp. 22-38. "Utilitarian generalization" is defined in terms that make it broadly comparable to rule-utilitarianism, as "the moral theory that an act is right or wrong depending on the value of those consequences which would obtain were everyone to perform acts of the kind being evaluated" (p. 22).

10. Ibid., p. 24.

11. David Collard, Altruism and Economy: A Study in Non-Selfish Economics (Oxford: Martin Robertson, 1978), pp. 30-35. Collard concludes that no system weighting in self-interest--which would include Gruzalski's sort of consequentialism--can solve the free-rider dilemma: only a Kantian system

of duty can.)

12. Gruzalski, "Defeat of Utilitarian Generalization", p. 31.

13. Ibid., p. 36.

14. Prior, "Consequences of Actions," p. 96.

15. D.D. Raphael, reply to Prior in same volume, pp. 118-119. Raphael thinks that Prior's attack on Moore as typical of the actual-consequences school of utilitarianism is incorrect, and that any points Prior makes should only be applied to Moore himself. He bases this allegation on Moore's preference for the phrase "best total results" over "best total consequences". This seems to me an unduly nitpicking criticism, and I shall refer to Moore in terms of consequences, as does Prior.

16. Williams, "Moral Luck", p. 38.

17. On this point, see Reichenbach, "On the Justification of Induction," Journal of Philosophy, v. 37 (1940), quoted in G.H. Von Wright, The Logical Problem of Induction (Oxford: Basil Blackwell, 1957), p. 208 fn.: "Belief can be the motive of action, but belief as such can never justify an action; only a justified belief can do that." See also Richard B. Brandt, A Theory of the Good and the Right (Oxford: Clarendon Press, 1979), p. 71.

18. Broad, "Doctrine of Consequences", p. 28.

19. Calculated from table two in S. Liechtenstein et.al., "Judged Frequency of Lethal Events," Journal of Experimental Psychology: Human Learning and Memory, v. 4 (1978), 556-7.

20. Charles Nesson, "Reasonable Doubt and Permissible Inference: The Value of Complexity," Harvard Law Review, v. 92, 1187-1225.

21. L. Jonathan Cohen, The Probable and the Provable (Oxford: Clarendon Press, 1977), p. 120.

22. J.J.C. Smart in Smart and Bernard Williams, Utilitarianism: For and Against (Cambridge: Cambridge University Press, 1973), p. 40.
23. Ibid., p. 41.
24. Bruno De Finetti, The Theory of Probability (Chichester: John Wiley, 1974), v. 1, p. 215, original emphasis.
25. Jeremy Bentham, Pauper Management Improved, quoted in Nussbaum, Fragility of Goodness, p. 89.
26. G.E.M. Anscombe, "Who Is Wronged?", The Oxford Review, 1967, no. 5, pp. 16-17; John M. Taurek, "Should the Numbers Count?", Philosophy and Public Affairs, v. 6, no. 4 (summer 1977), 293-316; Derek Parfit, "Innumerate Ethics", Philosophy and Public Affairs, v. 7, no. 4 (summer 1978), 285-300; and the correspondence between Parfit and Charles Fried in Philosophy and Public Affairs, v. 8, no. 4 (summer 1979), 393-397.
27. Anscombe, p. 17, original emphasis.
28. Parfit, "Innumerate Ethics", p. 300. The question of mistakes in moral mathematics is also relevant to Nagel. On p. 31 of "Moral Luck", he suggests that in cases of decision under uncertainty, "it seems rational to subtract the effects of occurrences subsequent to the choice that were merely possible at the time, and concentrate moral assessment on the actual decision in light of the probabilities." This is incoherent: the probabilities will only sum to one and be internally consistent if they represent the entire universe of possible outcomes.
29. Gruzalski, in "Defeat of Utilitarian Generalization", asserts that most recent literature on act and rule utilitarianisms uses the actual-consequences formulation.
30. This relates to the point about "getting it right" for the individual

defendant in the Nesson example versus "getting it right" for the group; but more broadly, see Jonathan Glover, Causing Death and Saving Lives (Harmondsworth: Penguin, 1977), chapter 16, section three, "Certain Death for a Known Person and 'Statistical' Death", pp. 210-213, and J.L. Mackie, Ethics: Inventing Right and Wrong (New York: Penguin, 1977), p. 196 ff.

Chapter Four

Rationality and Prudence

In my previous chapter I touched on the assertion that it must be possible to employ probabilistic reasoning in ethical choice because it is used in everyday prudential decision-making. This claim is made by Smart and many other utilitarians, normally coupled, as in Smart, with an assumption that everyday prudential decision-making can actually teach us something about ethics. But what if the twain shall never meet? Most of my attention in this chapter will be directed towards this claim that we should be able to use prudential reasoning as a model for ethics, and the related but broader supposition that ethical choice and rational or prudential decision-making are one and the same. That second position will in turn be subdivided, so that there will be several possible propositions about the relationship of prudence, rationality, and moral choice. Among them I discuss the following permutations:

1. Prudent decisions are moral decisions, though imprudent decisions are not necessarily irrational. This is the view which I attribute to Derek Parfit.

2. Rational decisions are the same as prudent decisions and also the same as moral decisions. This I take to be Richard Brandt's claim.

3. Irrational decisions are imprudent but not necessarily unethical; the prudent is not coterminous with the ethical. This is the Kantian position, and also my own; but it is also found in a different form in the work of Robert Nozick and David Richards.

That is, I want to ask whether or not there is an ethical realm which is separate from that of prudence, whether there is a class of decisions which we are not happy merely to get right in the long run, as we are with everyday consumer choices and similar prudential cases. It should be clear by now that a Kantian would reject the equation of the rational and the ethical, and equally obvious that I want to keep a firm boundary between the two, in the interests of retaining a notion of agency but limiting what agents are responsible for. This, it will be recalled, is my preferred way out of the moral luck paradox--or at least my method for reducing the paradox's impact. At the end of this chapter I also discuss briefly whether we should try to eliminate it altogether, in relation to Thomas Nagel's claims in The View from Nowhere.

However, before I begin the substantial discussion, I want briefly to query Smart's prior assumption, that we do use probabilities in ordinary prudential decision-making. At first this may seem an unexceptionable assumption. Indeed, the principle of prudence is often stated in terms of statistical aggregates, probabilities and utilities: "that one ought to

maximize the sum of intrinsic value or minimize the sum of intrinsic disvalue for oneself over one's lifetime as a whole."¹ But on the other hand, many psychologists and other students of actual decision-making have cast serious doubt on the claim that agents are "rational" enough to use probabilities.² Cognitive and motivational biases are said to dominate, so that there is no model of rational decision-making in everyday affairs which could serve as a parallel for moral choice. Perhaps we still ought to make everyday decisions on the basis of careful calibration in probabilities, but that is another claim, and a weaker one. To some extent, Smart does realize that the empirical basis of his first claim is dubious: "The fact that the ordinary man thinks that he can weigh up probabilities in making prudential decisions does not mean that there is really any sense in what he is doing."³ Granted, gladly; but then the direction changes. Smart is really trying to apply a moral imperative--that in favour of making decisions "rationally"--to everyday decision-making, rather than taking ordinary choices as a model for ethical ones.

What then is the relationship between ethical and prudential reasoning? In Butler, for example, there is a strong connection between them, and a further link between the prudential and the probabilistic. These pairs of relations are transitive, so that Gladstone could write: "Butler's fundamental contention is that probability involves moral obligation."⁴ But it is not always clear whether Butler's equation of the ethical, the probable, and the prudential is anything more than a semantic muddle. For example, he calls a probability of 1.00 that there will be frost on at least one day this winter "a moral certainty, i.e., ground for an

expectation without any doubt of it."⁵ Perhaps a better starting point is Derek Parfit's argument that the grossly imprudent is also immoral.

According to Parfit, the imprudent person is effectively treating his future self callously, as a different person. The ethical criterion of respect for persons then applies: he is obliged to treat his future self with whatever degree of respect he should evince for others. A young man's decision to take up smoking is not only imprudent but also unethical, because he is imposing on another the risk of a painful early death.

This boy does not identify with his future self. His attitude towards this future self is in some ways like his attitude towards other people. This analogy makes it easier to believe that his act is morally wrong. He runs the risk of imposing on himself a premature and painful death. We should claim that it is wrong to impose on anyone, including such a future self, the risk of such a death. More generally, we should claim that great imprudence is morally wrong. We ought not to do to our future selves what it would be wrong to do to other people.⁶

What is interesting, and atypical, is that Parfit refuses to call the teenager's decision to smoke irrational. In this he differs from Richard Brandt, for example, whom I shall discuss later in this chapter. In fact, it is the inappropriateness of calling an imprudent action irrational which compels Parfit to label it immoral. It cannot be termed irrational to be unconcerned about someone else's welfare--including that of my future self, Parfit says. But we still need some way of criticising that attitude, he feels. Let us therefore call it immoral, Parfit concludes.

But why not just call it a gamble? It would seem more straightforward simply to say that the young smoker has made a bad bet if he turns out to die of lung cancer or heart disease, and a profitable gamble if he does

not. If he "wins" his "bet", if he survives to his threescore and ten, would Parfit still want to term his decision unethical? The answer appears to be yes, and in this respect Parfit is actually taking a stringent Kantian position on moral luck.

What is wrong is not simply to kill another person, but to impose on him the risk of a racked, painful death. If the identity argument is granted, this seems a plausible position in Parfit's example, but perhaps only because the risks imposed by smoking are high. Cigarette smoking is the main cause of premature death in the United States.⁷ But would it be equally wrong to impose on oneself or another the risk of dying from a very minor cause of death? In chapter seven, on informed consent in medicine, I examine this question further.

In terms of Parfit's argument, there is a paradox. If Parfit wants to maintain that even minuscule risks should not be imposed, he will have to abandon the ordinary distinction between prudent and imprudent risks, good and bad gambles. Yet the procedure seemed to be equating the imprudent and the immoral, that is, knocking down the barriers between the prudential and ethical realms. That would imply that the smaller the risk, that is, the less imprudent the decision, then the less ethically wrong would be the imposition of the risk. And yet what was wrong was not to kill another person, but to impose a risk of death. That would indicate that it does not matter how unlikely something actually is to kill.

If it is imposing any risk which is morally wrong, the prudent will not be coterminous with the ethical. Smart's proposition will actually be reversed: however much we rely on cost-benefit reasoning in ordinary life, we will be forbidden to use it in moral decisions. That is, if an action carries with it a probability greater than zero of harm to one's future self or another person, it will be forbidden. Now in some cases this is clearly correct. Since I have a responsibility to provide for my dependent children, I would be morally in the wrong if I failed to insure my life, whether or not I die before they grow up. It would be taking risks with their security which would be wrong, whatever the actual outcome. In this case the imprudent and the immoral would be coterminous. But the upshot of Parfit's position would seem to be more extensive than that. Even if a risk is prudent, it would still be disallowed simply because it is a risk.⁸

This itself carries a risk: broadening the realm of the unethical beyond that of the imprudent, to an obsessive and paralysing extent.

In contrast, Robert Nozick seems to suggest that there are indeed duties not to take risks with other people's futures--leaving out the question of whether one's later self constitutes another person--but that they are limited by certain prior considerations.⁹ Even if Jane Doe knows that she will be fully compensated for enduring the risk of having her arm forcibly broken at some unspecified time during the next month, there is still something wrong in taking that kind of risk with her future. This has nothing to do with the level of the risk, but with the general legitimate interest in avoiding a nervous, fearful kind of society. But taking risks with other people's futures is not prohibited in blanket fashion. Doing so

would itself create risk and insecurity, since no one would be immune from the operations of moral luck. As long as an agent was deemed to have set the probability machine in motion, he or she could be punished for any harmful consequences, no matter how low their probability.

Exactly what kinds of risks would Nozick forbid us to take with other people's futures? Nozick does also consider how great a probability of social consternation would trigger off a ban. What must be considered is the cumulative total of acts, each of which taken individually may only involve a low probability of harm. (This is relevant to nuclear deterrence, and comes up again in chapter eight.) There is a case for banning the totality of such acts, but the law does not generally deal with groups of acts or agents¹⁰. To forbid each individual action, as a natural rights theory might, would also verge on a blanket prohibition on all risk-taking, and would tell us nothing about how great a probability of harm is necessary for rights to have been threatened. There is a sort of paradox of the heap here.

Each individual act's probability of causing harm falls below the threshold necessary for apprehension, but the combined totality of the acts may present a significant probability of harm. If different persons do each of the various acts in the totality, no one person is responsible for the resultant fear. Nor can any one person easily be held to cause a distinguishable part of the fear. One action alone would not cause fear at all due to the threshold, and one action less would probably not diminish the fear.¹¹

In the end, for such reasons, Nozick finds it difficult to apply probabilistic, prudential calculation to ethical decisions. The calculus of probability cannot itself reveal how great a probability of harm

constitutes unjustifiable imposition of risk on others. Any cut-off point will appear arbitrary.

If Nozick is less willing to equate the prudential and the ethical than was Parfit, Richard Brandt goes further in the other direction, that of collapsing the prudent, the right, and the rational into a single category of value. Brandt begins by asking two questions:

- 1) What would a person (perhaps all persons), if rational in the sense of having made optimal use of all available information, want and choose to do?
- 2) What kind of moral system, if any, would such a person support for a society in which he expected to live?¹²

The rational person would want and choose certain kinds of actions, and these actions are later identified with prudence¹³. The same rational attitudes which create prudent choices are also the touchstone for the ethical system. So it is clear that Brandt identifies the ethical, prudential, and rational decision-making attitudes as one. It is therefore incorrect to treat him as simply a theorist of prudence, if prudence is merely conceived as an adjunct to normative moral theory, even if an essential one.¹⁴ For Brandt, the theory of prudence, recast as a model of rational decision-making, is not a minor satellite of the ethical sun, but a guiding light in itself. This becomes even more apparent when Brandt asserts that all the traditional questions of ethics really represent questions about rational attitudes and behaviour.

[The methodological section of the argument] will assign a clear and useful meaning to the term "rational", explain why the term in that sense will naturally be commendatory, and argue that the... traditional questions of moral philosophy can be satisfactorily rephrased as questions about what it is rational to want for itself, rational for a person to do from his own point of view, and about which forms of action would be permitted by a moral system for a

society which it would be rational for a member of that society to support.¹⁵

Brandt presents his work as reacting against Rawls's modus operandi for ethical judgment, that is, to pretend that we know less than we actually do. Instead of imposing a veil of ignorance, Brandt advocates the application of cognitive psychotherapy and empirical psychology. This goes some way to getting round my initial objections to the argument that we can generalise from the application of prudential judgment in everyday decisions to create a model for ethical choice. Brandt could accept that we do not in fact demonstrate perfect calibration in ordinary choices, but his normative model is a refined version of prudential judgment, purged through the application of empirical psychology to both the values we seek to maximise and the method by which we seek to maximise them.

To Brandt a prime, if not necessarily the prime, question is:

...how far actions, desires and moral systems can be criticised by appeal to facts and observations...That question everyone will want to answer, although the precise form of the question, and the answer which can be given, will depend on the state of knowledge at the time when it is asked...With this question in mind, I shall pre-empt the term "rational" to refer to actions, desires or moral systems which survive maximal criticism and correction by facts and logic.¹⁶

Now I very much doubt that in fact "everyone will want to answer" this question as a matter of priority, and some theorists would not have wanted to answer it at all. Plato would have thought it quite misguided, for example, and Nietzsche would have found it distinctly boring. But having registered a disclaimer about the first part of Brandt's assertion, I want mainly to examine the ensuing second part, that "the answer which can be given will depend on the state of knowledge at the time." I shall try to

show that this is a variant of Smart's confidence in a technical fix which would finally place utilitarianism on an unshakeable foundation. This is of course an old concern: Sidgwick also thought that "a more scientific process" of assigning probabilities would strengthen utilitarianism. But Sidgwick thought that improvement would be limited by three factors: the extent to which we can make our past judgements consistent with each other, the barriers to our comprehension of others' past experience, and the limitations on how well we can assess what will make us happy in the future.¹⁷ Brandt's confidence seems greater: he does appear to believe that he can turn up rational value systems through his examination. Indeed, he remarks that his success in so doing has exceeded his own initial expectations.¹⁸

Brandt offers three senses of the term "rational". An action is rational to a first approximation if all relevantly available information is at the focus of the decision-maker's attention. The second variant is, roughly, the action which the agent would have chosen had he undergone cognitive psychotherapy whilst making the decision. Brandt's final definition of "rational" combines the first two: an action is rational if it is motivated by rational desires and aversions as per sense two, and if the conditions for the first approximation are met, in terms of availability of information. This first definition is meant to be the easiest to satisfy, although I shall point out in chapter six on secrecy and moral luck that its implementation is very contentious. But if it is not satisfied, the third and final criterion cannot be met. What then does Brandt have in mind by this first approximation?

The example Brandt gives is that of an academic who picks the "wrong" university for his sabbatical: he "overlooks the (possible) fact"¹⁹ that a second university which he had considered briefly actually possessed materials crucial to his research. Brandt's phrase "overlooks the (possible) fact" is confusing, although perhaps it translates as "possible, for all he knew." Whatever "the (possible) fact" means, it does seem clear that this example is a matter of retrospective rather than future judgement, that it does not involve predicting the outcome of one's decision as Gauduin and Anna Karenina have to do. In this case we are dealing with hindsight: the decision is already taken, and the consequences have become apparent. In other words, there is no uncertainty here, and that in itself makes it an odd example. Decisions under risk--which are meant to be elucidated by prudential models--always involve uncertainty.²⁰

Now of course one can be in a state of uncertainty (through ignorance) about the past as well as about the future. Conversely, prediction can be of things gone by: Darwinian theories were confirmed by the "predictive" power which they gave to paleontologists, to unearth evidence of unsuspected phrases of the past.²¹ However, none of this applies to Brandt's academic: he had the facts about both universities' libraries at his fingertips, but overlooked one set. Or perhaps it would be more correct to say that the relevant information on both universities was freely available to the academic, but that the facts about the second university were not at the focus of his attention because he chose to ignore them, having been seduced by the first university's better climate or more attractive ambiance. This seems not to be cognitive uncertainty at all, but

motivated bias, if anything, and perhaps a better illustration of irrationality under Brandt's second definition of rationality than the first.

Let us return to the point about hindsight, however. Recall that Williams was open to criticisms about hindsight insofar as he argued in "Moral Luck" that only success could have justified Gauguin's decision. But how could Gauguin have known in advance whether he would succeed? Perhaps generalising from the example of the academic, in which hindsight is not a problem, Brandt stakes a great deal on our being able to know the rational or best action at the time we make our decisions. Moral luck is not a quandary for him because he believes that we can control the outcome of our choices at least minimally, through foresight. Nor is he disturbed by Russell-style doubts about using actual consequences as the touchstone of right action, a scepticism based on the vulnerability of actual outcomes to external factors. Brandt appears willing to judge by actual rather than possible consequences in many of his examples:

It is a fact that people are uncomfortable if a decision they make turns out to have distressing consequences, when they know it would have been avoided by fuller or more careful reflection. For instance, if one buys a car which turns out to hold the road poorly and to consume large quantities of petrol, and if one knows one could have anticipated these facts by perusing an easily available copy of Consumer Reports, one is quite annoyed with oneself. Now a rational action is by definition one which avoids all mistakes deriving from inadequate reflection.²²

Nor does this confidence in our ultimate ability to get it right through the application of rational principles apply only to decisions conventionally regarded as prudential. Nussbaum has pointed out the complete absence of a tragic sense in Brandt's assertion elsewhere that "If

a person is disappointed in love, it is possible to adopt a vigorous plan of action which carries a good chance of acquainting him with someone he likes at least as well."²³ But of course this is consistent with Brandt's expansionist sort of rationalism, which claims for rational decision-making the power to handle all the traditional questions of moral philosophy. There are no insoluble dilemmas, no true conflicts of duties, no source of tragedy, on this view.

Still, at the very least Brandt should temper his confidence with the realisation that statistical independence means Consumer Reports' general advice is at best correct only in the long run. It cannot describe the particular car I am considering: only the class of all cars of that model and perhaps year. It gives a typical mileage figure for this statistical aggregate, derived, if done properly, by testing a statistically valid number of vehicles chosen at random. Even assuming that these procedures are followed to the letter, the report will only be correct in the long run. Nothing stops it being right in the long run if the particular car I have the misfortune to buy turns out to be the odd "lemon". But Brandt defined a rational action not as one which took Consumer Reports into account, but as one which avoids all mistakes deriving from inadequate reflection. Adequate reflection would have taken statistical independence into account. Brandt also said that the distressing consequences would actually have been avoided by fuller reflection. But the fuller reflection which encompasses the concept of statistical independence does not actually guarantee that I will not wind up with a lemon: it merely warns me that

this remains a possibility despite my having made the decision in the best possible way.

There is an element of the technical fix here. Brandt asserts that "It will hardly be questioned that one's choices should be affected by right estimates of the probability of certain outcomes."²⁴ Sometimes, however, even indubitably correct probabilities must actually be ignored in order to get the ethically correct decision. This was so in Nesson's example of the twenty-five prisoners, discussed in the previous chapter. I do not intend to repeat those arguments here, merely to point out that Brandt seems to be unaware of them.

Indeed, Brandt relies heavily on probability in what he calls the "law" of action-tendency:

If an individual expects, to a degree E, that a consequence of an action A by him now will bring about an outcome O, then, if O has a positive valence for him, his tendency (T) to perform $A(T_A)$ will have the magnitude of the product E times V. If O has a negative valence V, the tendency (T) not to do $A(T_A)$ will have the magnitude E times the absolute value of V.²⁵

Now this "law" is actually one of several optional strategies in decision theory, usually called expected-value or expectation strategy. It has no status as an empirical law of human behaviour: it is a normative principle for rational choice, and not one which is appropriate in all instances. For example, if probabilities are unknown, or if they are cancelled out by infinitely high or low utilities, as in Pascal's wager, another strategy is often thought appropriate: maximin, in which probabilities are deliberately ignored and the "least worst" option chosen. I argue in the the chapters on

nuclear deterrence and on allocation of scarce medical goods that maximin is indeed more appropriate in both these cases. Rawls also asserts that it would be used by those in the original position to rule out a society in which slavery would be sanctioned, for example. He has been criticised for this claim by those who argue that expectation strategy would be more appropriate, but even Rawls's critics would hardly term expectation strategy a universal law.²⁶ Yet another possible decision strategy is the principle of maximising conditional expected utility, which advises the agent to perform that act which makes the most desirable outcomes the most probable;²⁷ another is Richards's lottery principle, discussed later in this chapter.

It seems that Brandt is trying to use his version of rationality, embodied in expectation strategy, as both a normative principle and a descriptive law. Indeed, he presents this as a strength:

"Rational," in my sense, would be both a descriptive term and a recommending term. In the latter respect it would then be like "ought" and "good" and "best" reasons...although these are unlike "rational" in my usage in not having any definite descriptive meaning.²⁸

This is fraught with danger for a scientific project like Brandt's, as Nigel Howard has pointed out in another context:

We wish "rational behaviour" to be simultaneously the way people do behave, the way that logically they must behave, and finally the way they should behave. But these are extraordinary demands, not fulfilled by any other science.²⁹

Yet Brandt also appears not to believe that his "law" is fully statistical in its probability operations. He is not happy to sum probabilities in order to get the total likelihood if several different outcomes are involved, and he refuses to commit himself consistently to cardinal

measurement. At one point he doubts whether we can have a quantitative formulation of probabilities at all³⁰--the formulation on which expectation strategy relies. This seems to be a case of wanting to have it both ways, to enjoy the certainty of a scientific law without the bother of its rigour. As Howard puts it, with attractive cynicism:

Very often, when a quantitative model is built, the social scientist [or philosopher] will wave his hands and say, "The actual numbers used in this model do not matter. One has to apply this result in a general, qualitative kind of way." Surely this is inadmissible.³¹

To recapitulate, then, there is no gap in Brandt between the procedure for making more efficient prudential decisions and that for making right moral choices.

So we think of individuals first undergoing cognitive psychotherapy with respect to the desires and aversions possibly relevant to the present question, to make these rational; and then, with the rectified desires and aversions, we think of them acting to select or support a moral system on the basis of an ideally vivid representation of all relevant available information. The person making the decision...will be disabused of "unauthentic" desires, and desires based on false beliefs or on childhood deprivations or on improper generalization of experiences; he will not overlook outcomes or wrongly estimate their probability, and he will weight them in an undistorted holistic way, and so on.³²

Neither is there any tension between our being held responsible for outcomes, over which we have imperfect control, and the simultaneous requirement that we should not be held responsible for matters beyond our influence. That is, there is no problem of moral luck. This has been achieved at the cost of assumptions I find untenable, particularly whether we typically know as much as Brandt thinks we do, and whether, even if we could control outcomes in prudential matters minimally by forecasting them correctly, we might still face incommensurabilities of moral value in

ethical decisions. In other words, I doubt that this is a resolution of the moral luck paradox, since the concept of moral luck would be quite foreign to Brandt; it is an indirect assertion that the paradox does not exist.

I want to end with David Richards's more limited model of human rationality, which may be more acceptable than Brandt's excessively grand one. Like Brandt, Richards is concerned to explore the principles of rational decision-making; unlike Brandt, he confines those principles to prudential decisions. That is, he represents a third alternative to Parfit and Brandt. In Parfit, an imprudent decision is immoral, although it may be impossible to label it irrational. In Brandt, an irrational decision is both imprudent and immoral. Richards, more conventionally perhaps, terms irrational decisions imprudent but not necessarily unethical.

The actual principles of rational choice employed by Richards differ from Brandt's in at least two ways. First, they do not always yield a solution--or if they do, it is likely to be a "satisficing" rather than a "maximising" function.³³ "Satisficing" means settling for less than maximum utility, roughly speaking, for a "good enough" rather than an ideal solution. Second, even if they could be made to yield a definite answer, we should not always let them do so. This would be a form of technological determinism, the belief that because we can do something we ought to do it. But as Tom Stoppard says in The Dog it was that Died, "something can't be right because it's inevitable". To maintain otherwise is a form of the naturalistic fallacy.

Before invoking the decision procedures for rational choice--and for Richards, there may be more than one procedure, whereas Brandt accepted only one principle and elevated it to the status of a law--we must decide whether this is a matter of ethical or prudential interest. Now Richards may seem simply to be asserting rather than proving the distinction between the rational/prudential and the ethical here. He is aware of this danger:

At every point of my account, the primacy of the principles of morality to the principles of rationality was emphasised; the characterisation of these principles is, clearly, a theoretical requirement on the adequacy of my account of rationality, for the principles of morality importantly determine when the principles of rational choice apply.³⁴

That this distinction is, however, something more than mere intuition was shown by our reaction to Nesson's prisoners example. The rational choice appears to be convicting twenty-four prisoners, chosen at random. If we do that, we avoid the paradox of maintaining that the probability of each prisoner's guilt was 1.00--if we convict all twenty-five--or 0.00--if we convict none--though we are certain that the likelihood is .96. Yet the rational procedure is even more deeply repellent, for reasons which I discuss in chapter three, than convicting all twenty-five men in the certain knowledge that one of them is innocent. What appears the "moral" choice, or at least the one which coincides with the Anglo-American presumption of innocence and the demands of treating each case as individual, is to convict none of the men. Yet this is profoundly unreasonable, in light of the facts--about which there is no argument.³⁵

Just as the relation between the moral and rational principles is hierarchical in Richards, so there is an ordering to the rational principles themselves, one which emphasises the importance of probability

judgments. The first dictum which the agent should invoke is the lottery principle. The simplest example of this is a choice between lottery tickets offering different sets of probabilities for objects of value. The rational agent will a) prefer the lottery ticket which gives the higher probabilities of winning; b) where probabilities are exclusive, prefer the ticket which gives the higher probability of winning a more valuable prize; and c) where probabilities are equal, prefer the ticket which yields the more valuable objects. Generalising from this example, Richards enunciates the lottery principle as the first element of a strategy for advancing self-regarding ends, which he equates with the domain of practical reason although not with that of ethical choice.

Given a plan of action which achieves x , the object of a desire, with probability p ($0 \leq p \leq 1$), and y , the object of another desire, with probability p^1 (where $p^1 \leq 1-p$), and another plan which achieves x with probability p_1 and y with probability p_1^1 (where $p_1^1 \leq 1-p_1$), then that plan is to be chosen in which the probability of securing x is greater than or equal to the probability in the other plan and the probability of securing y is similarly greater than or equal to that of the other plan, except where $p=p_1$ and $p^1=p_1^1$, when the plans are to be indifferently chosen. Where plans of action secure the same ends (e.g. x and y , as above), and $p^1=1-p_1$ and $p_1^1=1-p_1$ (i.e. the probabilities being exclusive), then that plan is to be chosen which gives a higher probability to the more preferred alternative. And if there are the same probabilities in two plans of action, which secure entirely different ends, that plan is to be chosen which secures ends, at least one of which is preferred (the others being indifferently preferred) to one of those secured by the other plan, all ceteris paribus.³⁶

If the lottery principle does not yield an answer, the decision procedure then directs the agent's attention to the following other principles, in order:

- 1) Dominance: given several plans of action, choose the plan which secures all the desired ends of the other plans, and more.
- 2) Effective means: given a desired end, choose the action which most effectively and at least cost attains that end.
- 3) Postponement: if it is unclear at the time of choice what the desired ends will be or how they will best be secured, put off making a choice among the plans of action which secure the desired

end.

4) Satisficing: if the other principles do not apply, choose that plan which, after a dispassionate, well-informed weighing of probabilities and preference intensities, will best secure one's desired ends.³⁷

Richards's procedures for rational choice appear more complete than Brandt's, more sophisticated in their treatment of probability, and clearer in their hierarchical ordering and applicability. Why does he make lesser claims for them than Brandt does for his principles of rationality?

Richards chooses to limit his principles to prudential decisions because the calculations become infinitely more complicated in moral choices. There, he says, we typically have to take into account our responsibilities to others, whereas prudential cases are usually concerned mainly with our own desires. This is no hard and fast rule, admittedly, but the second proposition is at least true of Brandt's two prudential cases. The examples of the professor choosing a sabbatical location and the driver choosing a new car do not incorporate the needs of others. How can Brandt simply generalise from them to moral choices? And in particular, what help would they be in medical ethics and political decisions?--areas in which responsibility to patients and accountability to democracies' populations are crucial.

Granted, Brandt also builds in cognitive psychotherapy, but his purpose there is simply to rid the decision-maker of "unauthentic" desires, wishes "based on false beliefs" or on "improper generalisation of experience." This "generalisation of experience" is not the Kantian requirement of universalisability which leads to the Categorical Imperative

and to respect for others. Nor is a desire to advance one's own interests, first and foremost, necessarily "unauthentic." The agent who has undergone cognitive psychotherapy "will not overlook outcomes", but there is no mention of whether or not he will overlook the claims of others. Brandt does not view failure to consider the wishes of others as an ailment to be cured by cognitive therapy, and indeed it is dubious that he could. As Parfit points out, it is not irrational to be self-centred.

Here Parfit would agree with Richards by implication: it may not be irrational to ignore the claims of others, but it cannot be ethical. (Therefore, to Parfit at least, if one's future self is conceived as an other, it is unethical though not irrational to impose risks on oneself in old age.) But actually this is a rather Kantian way of looking at things. By contrast, utilitarianism, according to Richards, "commits the profound conceptual error of assimilating moral principle to individual rational choice. It supposes that what is legitimate in the case of rational choice, namely, to satisfy certain desires less in order to satisfy certain desires more, or to refuse to satisfy certain desires because of wishing to satisfy other desires later, is legitimate in the case of the moral choice of satisfying the desires of different persons."³⁸

Richards does not propose a complete split between the ethical and the prudential realms; nor does he deny that some aspects of prudential reasoning may be instructive in ethical choices. Using a proto-Rawlsian device which he terms "moral amnesia" in preference to Rawls's "veil of ignorance", Richards asks what ethical standards would be accepted by

perfectly rational egoistic people, from a position of equal liberty, in uncertainty about their own desires, nature and circumstances. Very roughly, Richards can be understood as claiming that practical rationality, which he equates with prudence, is necessary for ethical choice, but insufficient to determine it. Other background conditions must also be specified. Nor is the project of making deliberation fully rational sufficient to motivate the ethical agent, as Brandt seems to think. Indeed, E.J. Bond has suggested that there is a paradox about rationality and motivation.

If practical rationality and morality are a matter of reason or cognition, which are objective and universal, they must lack the power to motivate, since that power depends upon the presence of the relevant contingent desires. If, on the other hand, practical (including moral) reasoning is confined within the limits of the agent's contingent desires, there can be no universal or objective reasons for action.³⁹

To posit a complete split between the ethical and the rational/prudential (to collapse these two together) seems untenable; but to equate them, as Brandt does, seems jejune. Furthermore, it is inadvisable from the point of view of moral luck. If we want to retain a notion of moral agency, we must limit what agents are responsible for as moral agents. It therefore seems undesirable to expand what counts as immoral to include crimes against one's future self, as Parfit does. But there is some link between the ethical and the rational: it is certainly true that irrational people are not expected to bear full moral responsibility for their decisions. This point will surface again in chapter seven, on informed consent in medicine. But the relationship is not

two-way: we do not simply say that evil people are irrational. Indeed, conventionally they are presented as cunning.

I have reached the end of the theoretical section of my dissertation. The dilemma of moral luck has been identified as arising from the incompatible requirements that we be held responsible for our actions and character partly according to the effects which they produce, and that no concept of responsible agency should hold people responsible for that which they cannot control. I have identified the problem as arising through asking Kantian questions, but have also suggested that Kant proposes some solutions rather than simply hoping the problem will go away. These escape routes are built around the notion of limiting what we are responsible for, as, for example, by exempting acts done under compulsion or by defining what counts as ethical more narrowly than is done by those theorists who equate rationality and/or prudence and ethics. But I have also suggested that consequentialism cannot ignore the problem of moral luck either.

In fact what I have been doing is to depict moral luck not as a Kantian hobby-horse, but as a dilemma which cuts across the boundaries of ethical schools, a general problem. Before leaving the theoretical section of the dissertation, I want to point out the parallel with what Thomas Nagel identifies in The View from Nowhere as the single problem which underlies philosophy of mind and ethics, what he calls the "problem of excess objectivity".⁴⁰ Nagel's concern is how to reconcile being in the world and being a person. "[H]ow to combine the perspective of a particular person inside this world with an objective view of that same world, the

person and his viewpoint included...is the most fundamental issue about morality, knowledge, freedom, the self, and the relation of mind to the physical world."⁴¹ This is very close to the tension between being responsible for our actions' often unknowable and generally largely uncontrollable consequences, and being responsible at all, being moral agents in any meaningful sense.

Nagel attributes "excess objectivity" to impersonal theories such as consequentialism, which he describes as "holding that we should try so far as possible to transform ourselves into instruments for the pursuit of the general good, objectively conceived (though our own interests play their part along with everyone else's in defining that good)."⁴² But he also reserves some ammunition for Kantianism, which he accuses of trying to ignore perplexities about objectivity and subjectivity. Now I noted before, in relation to Nussbaum and moral luck, that simply because Kant proposes a solution does not mean that he ignores the problem. And there is a risk of excess romanticism in holding with Nagel that "Certain forms of perplexity...embody more insight than any of the supposed solutions to those problems."⁴³ This is a particular dilemma, in practical ethics, although on a personal note, I have found when teaching medical ethics to nurses that health professionals do not expect philosophy to give them answers: merely to teach them to ask the right questions.

In this later work Nagel's scepticism is less extensive than Williams's in Ethics and the Limits of Philosophy, just as in the original moral luck debate; but it is nonetheless substantial, particularly where

the concept of agent-responsibility is concerned. I have hoped to retain a sense of responsible agency by limiting what agents are responsible for, and in the practical ethics chapters, I shall continue to follow this line, which seems suitably pragmatic. But I do want to say something on the subject of Nagel's doubts about the concept of moral responsibility, and how he thinks we might learn to live with them, to retain a notion of agency, although a much revised one.

Nagel thinks the problem of moral responsibility is unsolved, and possibly insoluble. He begins by setting up a dichotomy very similar to the one about moral luck, though couched in terms of objectivity and subjectivity:

Once people are seen as parts of the world, determined or not, there seems no way to assign responsibility to them for what they do. Everything about them, including finally their actions themselves, seems to blend in with the surroundings over which they have no control. And when we then go back to consider actions from the internal point of view, we cannot on close scrutiny make sense of the idea that what people do depends ultimately on them. Yet we continue to compare what they do with the alternatives they reject, and to praise or condemn them for it.⁴⁴

What Nagel terms "the radically external standpoint that produces the philosophical problem of responsibility"⁴⁵ seems to erase agency altogether. This criticism was first raised in relation to extreme consequentialism, but Nagel also means it to apply to Kant. However, reading Kant in terms of freedom rather than only responsibility gives a result which Nagel should find more conducive. Nagel asserts that "To be really free we would have to act from a standpoint completely outside ourselves, choosing everything about ourselves, including all our principles of choice--creating ourselves from nothing, so to speak."⁴⁶ But

something very much like this is going on with the determination of the will directly by the Categorical Imperative. Indeed, later Nagel identifies himself as agreeing with Kant's view that ethics and freedom are linked⁴⁷.

Furthermore, although Nagel says the problem of responsibility has no solution, he does think that we can live with it, and that we can confer a kind of freedom on ourselves. This can be done through "a kind of reconciliation between the objective standpoint and the inner perspective of agency which reduces the radical detachment produced by initial contemplation of ourselves as creatures in the world."⁴⁸ The maxim which Nagel produces is this: "Since we can't act in light of everything about ourselves, the best we can do is to try to live in a way that wouldn't have to be revised in light of anything more that could be known about us."⁴⁹ This has something to do with avoiding remorse but being able to express regret. It does not mean "never having to say I'm sorry", but perhaps being able to say simultaneously "I'm sorry about how things actually turned out" and "ich könnte nichts anders," to paraphrase Luther. Specifically, in relation to the moral luck argument, it is a refusal to say with Williams that we can manage nothing better than judgements from hindsight, which cannot be ethical judgements at all.

FOOTNOTES

1. Joyce Trebilcot, "Aprudentialism", American Philosophical Quarterly, v. 11, no. 3 (July 1974), p. 204.

2. See, for example, S. Liechtenstein et.al., "Judged Frequency of Lethal Events"; D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgement under Uncertainty (New York: Cambridge University Press, 1982); and George Robinson and Avraham Merar, "Informed Consent: Recall by Patients Tested Postoperatively", in Samuel Gorovitz et.al., Moral Problems in Medicine (Englewood Cliffs, New Jersey: Prentice-Hall, 1983, second edition), p. 182 ff. The question of patients' rationality is taken up again in chapter seven on informed consent.
3. Smart, in Smart and Williams, Utilitarianism: For and Against, pp. 40-41.
4. Joseph Butler, The Analogy of Religion, ed. William Gladstone (Oxford: Clarendon Press, 1896), footnote to section four (which contains the famous quotation, "But to us, probability is the very guide of life.")
5. Ibid., section three.
6. Derek Parfit, Reasons and Persons, (Oxford: Oxford University Press, 1984), p. 320.
7. John H. Holbrook et.al., "Cigarette Smoking and Cardiovascular Diseases", updated version of 1977 publication, American Heart Association Report of Ad Hoc Committee on Cigarette Smoking and Cardiovascular Diseases for Health Professionals, reproduced in Tom Heller et.al., Coronary Heart Disease: Reducing the Risk, A Reader (Chichester: Wiley, 1987), p. 118.
8. I do not mean to imply that Parfit himself takes this strong a stand, but I do think it is a possible implication. Parfit's reliance on the smoking example, in which the risks are high, may mean that he fails to see the corollary problem about imposing even low risks.
9. Robert Nozick, "Prohibition, Compensation and Risk", in Anarchy, State

and Utopia (New York: Basic Books, 1974).

10. Partial exceptions would be class-action suits in U.S. law, and perhaps the law of conspiracy.

11. Nozick, p. 73.

12. Richard B. Brandt, A Theory of the Good and the Right (Oxford: Clarendon, 1979), p. v.

13. Ibid., p. 81.

14. As is done by Richard M. Hare, in "Ethical Theory and Utilitarianism", in Amartya Sen and Bernard Williams (eds.), Utilitarianism and Beyond (Cambridge: Cambridge University Press, 1982), at footnote 10, p. 27.

15. Brandt, Good and Right, pp. 1-2.

16. Ibid., p. 10. Perhaps Brandt exaggerates the difference between himself and Rawls, who has written: "...[A]dopting the role of observing moral theorists, we investigate what principles people would acknowledge and accept the consequences of when they have had an opportunity to consider other plausible conceptions and to assess their supporting grounds. Taking this process to the limit, one seeks the conception, or plurality of conceptions, that would survive the rational consideration of all feasible conceptions and all reasonable arguments for them." Rawls characterises this procedure as "a kind of psychology". ("The Independence of Moral Theory", Proceedings and Addresses of the American Philosophical Association, v. XLVIII, 1974-5, p. 8.)

17. Henry Sidgwick, The Methods of Ethics (London: Macmillan, 1962, 7th edition), p. 142.

18. Brandt, Good and Right, p. v.

19. Ibid., p. 11.

20. I am using "risk" and "uncertainty" in the common manner, rather than according to the stricter usage in decision theory, which distinguishes the two. In the narrow interpretation, a risk decision is one in which the outcomes and probabilities of each outcome are known; a decision under uncertainty is one in which the outcomes are known but the probabilities are not. See D. Ellsberg, "Risk, Ambiguity and the Savage Axioms", Quarterly Journal of Economics, v. 75, pp. 643-669.
21. Stephen Toulmin, Foresight and Understanding (London: Hutchinson, 1961), p. 23 ff.
22. Brandt, Good and Right, p. 153.
23. Brandt, "The Morality and Rationality of Suicide", in J. Rachels (ed.), Moral Problems (New York, 1975), pp. 363-387.
24. Brandt, Good and Right, p. 154.
25. Ibid., p. 48.
26. For example, John C. Harsanyi considers that Rawls is technically wrong to use maximin, and even that maximin is an irrational decision principle. But he still recognises it as a strategy, although dismissing it as an outmoded one. ("Morality and the Theory of Rational Behaviour", in Sen and Williams, Utilitarianism and Beyond, p. 47.)
27. See Ellery Eells, Rational Decision and Causality (Cambridge: Cambridge University Press, 1982).
28. Brandt, Good and Right, p. 152.
29. Nigel Howard, Paradoxes of Rationality (Cambridge, Massachusetts: M.I.T. Press, 1971), p. 49.
30. Brandt, Good and Right, p. 75.
31. Howard, Paradoxes of Rationality, pp. 1-2.

32. Brandt, Good and Right, p. 202.

33. This distinction first occurs in Herbert Simon's Models of Man (New York: John Wiley, 1957). It has been picked up again by Michael Slote in his concept of "satisficing consequentialism": see footnote 16, chapter eight.

34. David A.J. Richards, A Theory of Reasons for Action (Oxford: Clarendon Press, 1971), p. 71.

35. The jury in the Nesson example is in a position to exercise rationality in Brandt's first sense of the term, the possession of all available information. On page 13, Brandt defines this as "the propositions accepted by the science of the agent's day, plus factual propositions, justified by publicly accessible evidence (including testimony of others about themselves) and the principles of logic."

36. Richards, Reasons for Action, p. 28.

37. Ibid., p. 43.

38. Ibid., p. 87.

39. E.J. Bond, Reason and Value (Cambridge: Cambridge University Press, 1983), pp. 6-7.

40. Nagel, View from Nowhere, p. 8.

41. Ibid., p. 3.

42. Ibid., p. 8.

43. Ibid., p. 4.

44. Ibid., p. 120.

45. Ibid., p. 122.

46. Ibid., p. 118.

47. Ibid., p. 135.

48. Ibid., p. 126.

49. Ibid., p. 127.

Chapter Five

Moral Luck and the Allocation of Scarce Medical Goods

Does moral luck apply to the allocation of dialysis machines? To the provision of scarce beds in an intensive care unit for children with heart conditions? To organ transplants? I want to show that it does; but conversely, I think that we can apply findings from this area back to the general paradox. That is, there is a two-way relationship between practical and theoretical ethics here.

Ethics has generally been willing to treat questions about the allocation of scarce medical resources as part of its province insofar as they constitute questions of distributive justice.¹ Alternatively, the debate here might seem to relate to general questions about how to make the best decision under conditions of uncertainty. In either case, questions about medical resource allocation might appear to have no specific relevance to moral luck, and this thesis would appear to divide into two separate sections: the first on moral luck, the second on decision-making under uncertainty, as illustrated in some spheres of medical ethics and practical politics. Further, decision-making under uncertainty might not even be an ethical matter, according to what boundary line is set up between rational consideration, prudent action, and moral behaviour--as discussed in chapter four. (One possible corollary of this second perspective is that allocating the kidney to the patient

with the highest probability of cure can be presented as a value-free decision. Some philosophers and health care professionals have taken this line, though I think fewer accept it these days. I certainly do not, and the reasons why not will become clear as the argument progresses, I hope.)

Of course it is important to consider decision-making under uncertainty and morally optimal outcomes, and if I can contribute any principles of choice I will reckon I have achieved something important. But this is not all I want to do. These practical chapters still retain the focus on responsibility and agency which characterised the theoretical half of the dissertation. It is crucial to remember who the agent is in cases of scarce resource allocation, however, or confusion may ensue. The patient is not the agent, and if ill-luck strikes her, if she does not get the kidney or the operation, that really is just ill-luck, no matter how serious the outcome. It is not moral ill-luck. The agent is the health care professional, and the moral luck questions which the professional faces are those about whether the agent's responsibility is to get the optimal distribution in light of available resources or to make the decision in a manner which maximises patient autonomy and demonstrates respect for persons; and whether the professional's duty is to the individual patient or to a class of sufferers. Broadly, I want to ask whether Nagel's maxim, which I identified in the previous chapter as related to the moral luck paradox, can be put into practice in making choices about scarce resource allocation in health care. Is it possible for the health care professional to act in a way

that would not have to be revised in the light of outcomes? And what if both choices available to the professional are evil? It is no coincidence that decisions about the allocation of operations and organs have been labelled "tragic choices."²

In the discussion of moral luck and scarce resource allocation which follows, I shall need to consider two other subsidiary questions. First, let me return to the matter of whether cases about scarce resource allocation are prudential or ethical matters, or perhaps both. Now there is something very strange about depicting choices which may result in life for one and death for another as merely prudential. Perhaps we ought to say that decisions which we must get right this time are ethical; those which we are content to get right in the statistical aggregate may count as prudential. It would appear obvious that life-and-death decisions have to come out right this time. This would be consistent with the impetus in the twenty-five prisoners example, in which it was not good enough to get an overall figure of twenty-four convictions and one acquittal by any means. Each man was being tried on his own, and each deserved to have his case kept separate from any considerations about the aggregate, if we wanted to maintain basic legal principles. We had to get each trial right; in each prisoner's case, we had to get it right this time.

But is this so? To begin with, what made it wrong to breach the ethical/legal principle of prior innocence in the Nesson case was not that this was a life-and-death example. It would still have been

wrong even if the death penalty were not in operation. And in fact it was not initially obvious what "getting it right" could mean in that case. The criterion in the Nesson example was not getting the mathematical conundrum right in each case, or even for the aggregate, but maintaining the principle of prior innocence. We were unable to "get it right" in mathematical terms: the only solution appeared to be acquitting all twenty-five prisoners, even though we knew that twenty-four were guilty. So "getting it right" this time may mean respecting personal autonomy or some other ethical criterion which occupies a similar rank to the principle of innocent until proven guilty. It may entail making the decision in a particular way, not necessarily getting the optimal outcome. But, as in the Nesson example, there remains a tension here between individual and group: these life-and-death decisions about resource allocation are made for classes of patients and for the individual patient to whom the medical professional owes a duty of care. And the agents in organ and dialysis allocation make decisions which are affected by the luck of the draw--so that prognoses made about statistical aggregates may not be true in this particular case.

Second, what will be the appropriate reaction for the health care professional who "gets it wrong"? That will require some consideration of remorse, regret, and Williams's additional concept, "agent-regret". This is actually linked to the question of whether questions of scarce medical resource allocation are prudential or ethical. We seem to want agents in decisions about allocation of scarce medical goods to be able to feel remorse, which does not apply

in prudential decisions which go wrong. There--in finding that my new car has turned out to be a lemon even though I followed "rational" decision procedures such as reading Consumer Reports--I may well feel regretfully that "I could kick myself," but I would hardly expect to feel guilty or remorseful. But I might expect remorse from a health professional who "wasted" a kidney on a patient who belied her good prognosis and conformity to the professional's predictive model by promptly dying. Guilt would again be too strong a feeling to demand: but somehow the health professional should feel differently, and more deeply, about the kidney case than I would about the car.

Would it help to introduce a third concept here, that of "agent-regret"? Williams connects this sentiment to failure not only of the agent's "project" but, more fundamentally, to his or her own failure.

It is felt specifically about the agent's past actions, rather than about external states which have gone wrong. That is, the agent is not divorcing herself from the success or failure of her project: its success may or may not be hers, but the actions which she took to set it in train certainly are. As far as Gauguin is concerned, says

Williams, "what would prove him wrong in his project would not just be that it failed, but that he failed."³ Gauguin's self-respect is

so intimately bound up in his art that the miscarriage of his artistic ambitions would have meant not only that his decision to abandon his commitments in Paris for Tahiti was wrong, but also that he himself was evil, to put the argument in a deliberately forceful fashion. I do not now want to discuss the difficult points about hindsight which I have treated elsewhere. The only point which I want

to resurrect here is the tightness of the link between Gauguin's sense of his own agency and worth, and his profession.

Clearly there is a parallel with health professionals here.

"Getting it wrong"--and I do not yet want to spell out what "getting it wrong" means in the case of kidney transplants--is more than just "project" failure to the health professional. The duty of care which the doctor or nurse owes to each individual patient forbids the agent in this case to write off a "failure" in such comforting terms as "better luck next time." This is the source of remorse: the combination of a lowered sense of one's professional self-esteem and failure to provide the best standard of care.⁴

But there is also a difference, and one which makes agent-regret less useful in medical ethics than Williams finds it in the Gauguin example. The medical professional owes a duty of care to the patient; it is hard to see how Gauguin owes a "duty to his art" except as a cliché. This makes it less useful to separate the outcomes for the patient of the health professional's actions from the actions themselves. What affects the patient is the outcome, not the intention or the action of the professional. So agent-regret, whilst perhaps a useful concept in other spheres, does not do justice to the relationship between health professional and patient.

Yet a health professional can hardly be expected to drag herself through agonies of remorse each time the statistically inevitable occurs, when the one patient in one hundred dies under a course of

treatment which has a probability of cure equalling 0.99. Nor would we expect the doctor to forgo that course of treatment for future patients. Indeed, if 0.99 is the best rate of success available from alternative treatments, a consequentialist, at least, would claim the medical professional was at fault if he did not prescribe that course.

Perhaps in such cases all we can ask the medical professional to feel, after all, is ordinary regret? As Williams notes,

Regret necessarily involves a wish that things had been otherwise, for instance that one had not had to act as one did. But it does not necessarily involve the wish, all things taken together, that one had acted otherwise.⁵

This is in fact surprisingly close to Nagel's maxim, that the best we can hope for is to arrive at decisions which would not have to be revised in light of our subsequent knowledge of how things turned out. Furthermore, Williams says,

While [the agent's] justification is in some ways a matter of luck, it is not equally a matter of all kinds of luck. It matters how intrinsic the cause of failure is to the project itself.⁶

It also matters how intrinsic the duty to "get it right this time" is to one's professional worth. Thus health care administrators and political decision-makers may not be bound by as strict a duty of care as medical professionals, although we would still want to say that they should feel regret over callous decisions or cases of corruption and bribery.

What other kinds of regret administrators and politicians should feel about health is beyond the scope of this chapter, although some

related issues about public versus private morality are treated in my chapters on secrecy and deterrence. The compelling cases are those of mixed feelings--standing by a "rational" decision which happened to turn out badly, or ruining a lucky but ill-made one. These cases abound in the area of scarce medical resources, made more compelling still by the existence of others who might be said to have a grievance--the patient who does not get the "wasted" kidney, the relatives of the patient who did get the kidney but died after "false" hopes had been raised, the old women in need of hip replacements who must wait years for a safe and cheap operation when medical resources are diverted to headline-grabbing transplants with a low rate of success. Indeed, it is the claims of others which can be argued to make this a moral rather than a merely prudential area.

It may help to make health professionals' duties less all-encompassing if we make a distinction between deciding how much of a scarce health resource will be produced and deciding how it will be allocated.⁷ We might want to claim that nature decides how many kidneys will be available, that nature also decides the prevalence of kidney failure, and that health professionals' responsibilities are second-order. Medical personnel do not set the probability machine rolling, after all, and to the extent that it rolls of its own accord, "deciding how much of a scarce health resource will be produced" misrepresents the kidney case. Of course this assumes a distinction between positive and negative responsibility which is open to debate: Helga Kuhse, for example, denies that letting die is different from killing, as does John Harris.⁸ In addition, deciding

how much of a health "good" will be produced can be a human function--the province of health ministers and the powerful medical interests whom they may consult--when treatments other than organ transplants are under consideration. Even in the case of transplants, human agency affects how many kidneys are available, since donors decide that--sometimes under the influence of other persons' actions. The carrying of kidney donor cards dropped dramatically after a television Panorama report alleging that organs were being taken from donors before brain stem death.

To return to the first subsidiary query: can the allocation of scarce health goods be made in a purely prudential manner? To come back to the second: can this be done so as to occasion no remorse in the health care professional? Why not decide on purely medical grounds, for example? Some doctors, particularly in Britain, still maintain that this is perfectly feasible, and if that were true, moral luck would not enter into it. But what constitutes the most medically correct choice is itself ambivalent. The most "savable" in terms of prognosis is unlikely to be the neediest or "illest" in terms of diagnosis. Nor could moral controversy be eliminated by the adoption of efficiency criteria in, say, a policy of triage, even among medical personnel--let alone between the health care professions and the general public. Indeed, the classic example of triage embodied conflict over ethics within the medical profession: between the medical officer in charge of the US North African forces in the Second World War and a consultant surgeon over whether scarce penicillin should be given to venereal disease sufferers rather than

to men wounded in battle--on the efficiency grounds that the VD sufferers could be more quickly returned to active service. Medical advisability is not a straightforward criterion on which agreement is possible, nor one which can eliminate tragedy and chance.

This general point can be made more specific with reference to the British model of allocation as described by Calabresi and Bobbitt.⁹ This clinical judgement model interprets the Aristotelian dictum on justice by treating as likes those potential recipients "who are equal with respect to exterior, observable, therapy-related criteria."

The criteria are applied unswervingly and damn the implications for general equality. Thus hemodialysis is allocated so as to achieve the highest rate of success, given a limited number of kidneys available.¹⁰

Now it is open to question whether the efficiency approach is self-defeating or self-contradicting even in its own terms. If there is a policy of not even considering those over a certain age as potential recipients, for example, how can we know what the success rate would be if they were treated? Is the highest level of possible efficiency actually being achieved? And how can we know whether this particular means of allocation meets the rational requirements of successful prudential thinking, in Brandt's terms? We can be even more sceptical about whether it fulfills ethical criteria, and still more so about whether it can avoid being judged on whether or not it is ethical.

This is clear at the most basic level from the outraged condemnation of the English system as profoundly unethical which Calabresi and Bobbitt found in Italian doctors whom they interviewed. In these

doctors' view, it is a sort of bad faith to claim that a purely clinical approach can avoid the operations of moral luck.

The Italian method of kidney allocation, as described by Calabresi and Bobbitt, confronts the need to take an ethical stance in dealing with scarce medical resources. The moral creed which it seeks to reflect is absolute egalitarianism, and favourable prognosis is abandoned as a criterion. This leads to what many would regard as a result which allows the vagaries of chance full play: under the Italian first-come-first-served system, a ninety-year-old patient with terminal cancer gets the kidney if he comes before the thirty-year-old with no other clinical symptoms. Surely this counts as failure and occasions regret, or even remorse?

Not if clinical merit is itself regarded as conferred by chance, the answer runs. Patients do not choose to suffer from terminal illness. "Why, after all, should their shorter lives be measured against lives that would have been longer from no merit of their own?"¹¹ What the Italian model does is to confront luck head-on and to draw its fire by incorporating it. (The reasoning parallels Rawls's treatment of intelligence--normally considered as a personal merit--as the gift of the Fates.) That the possibility of failure--in terms of inefficient use of resources--has been taken into account is clear from interviews with doctors who reiterate their loyalty to the "principle of humane non-interference even if, as a result, lives were lost."¹² There can be regrets in that case, and a certain amount of hedging: the doctors normally reserve one or two emergency

dialysis machines for the hard cases. Beyond this mild hypocrisy, however, the Italian model is considerably more honest than the British, which fails to confront the medically inescapable fact that members of the middle class have a better prognosis of recovery, since they are educated enough to understand the importance of the necessary diet and affluent enough to afford it.¹³ Thus ethical issues of justice and fairness, once again, creep into supposedly clinical judgments.

Nevertheless even Italian egalitarianism does not produce complete fairness--not even on its own terms. If "first come first served" operates as a kind of lottery at an individual dialysis centre, it does not do so for the country as a whole. On a per capita basis, there is a sizeable distinction between dialysis places available in the industrially developed north and in the poorer south.¹⁴ And of course first-come-first-served may not even be fully random at an individual centre: perhaps middle-class people will be more assertive or more aware of the centre's existence, and thus more likely to turn up for treatment.

But it seems fair to ask, as Williams might, whether such ill-chance is external to the doctors' own project of treating all those who come to them on a basis of strict equality. It may constitute project failure or perhaps even personal failure for national health administrators or politicians, but not personal failure for an individual doctor, whose duty of care is to the individual patients or potential patients who come before him. Such questions of public

policy are beyond the scope of this chapter, although I recognise that this omission leaves open important questions of bias towards wealthy regions, high-tech acute care, and predominantly middle-aged male diseases such as coronary conditions. But these questions are not so amenable to consideration in terms of moral luck. As a matter of individual agents' responsibility, I do not see how the failure to achieve countrywide fairness can occasion remorse in the individual physician.

The key paradox here is that only by openly recognising the operations of chance can one avoid vulnerability to moral luck. The second-order decision to randomise treatment, either through a lottery or, slightly less thoroughly, through a policy of first-come-first-served, remains a matter of ethics, not prudence, and is presented proudly as such by the Italian doctors. But I shall delve into public policy just deep enough to point out at the first-order level of deciding how much of a given health good is to be produced, lotteries may be counter-productive. As Calabresi and Bobbitt note, "lotteries tend to attract attention."¹⁵ We do not care to be reminded that luck is so endemic to the human condition that our lives may depend on it. If it becomes known that hospitals are allocating life and death on a random basis, through a lottery or a first-come-first-served procedure, fewer people may bequeath their kidneys, feeling that their gift of potential life is not being treated with due gravity and gratitude. This likelihood is mitigated, in a cynical view, by the certainty that power-holders would do their damndest to eliminate scarcity of medical resources if they had to

take their chances in a lottery, rather than being favoured by decision procedures based on "social merit."¹⁶

Lotteries may be dynamite, but less so than the Seattle "God" Committee's standards of "social merit", including income, net worth, church membership and Scout leadership.¹⁷ Not surprisingly, the Committee's recommendations turned out to produce what Calabresi and Bobbitt tactfully term "a pattern of preferences." It was noted that "the Pacific Northwest is no place for a Henry David Thoreau with bad kidneys."¹⁸ To their credit, the members of the committee eventually recognised that they were producing biased judgments.¹⁹ The consequence was regret so massive as to close the committee down.

If clinical criteria were unable to remain purely within the realm of prudence, still less, of course, can social criteria for allocation of scarce medical resources escape being ethical judgements. And in legal-ethical terms, lotteries have been held fairer than social criteria when there is conflict of life with life. In U.S. v. Holmes (1841), the presiding judge ruled that a surviving crew member, Holmes, should not have collaborated with his mates in devising and implementing social criteria for deciding who among a shipwreck's survivors must be thrown off a lifeboat in order to lighten its load. Despite his counsel's contention that the crew's method of selection--"not to part man and wife, and not to throw over any woman"--was more humane than a lottery, Holmes was convicted of unlawful homicide. In the judge's opinion, only casting lots would have been a remedy which the law could sanction: "In no other way

than this or some like way are those having equal rights put on an equal footing, and in no other way is it possible to guard against partiality and oppression, violence and conflict."²⁰

Holmes is a particularly telling case because it was deeply permeated by moral luck. The crew failed to prevent female deaths: two sisters jumped overboard to drown with their brother, who was among the fourteen men jettisoned. Now of course the justice of drawing straws could also have been threatened by the unexpected: a husband or wife who drew a long straw might well have chosen to die with an unlucky partner. But would this be cause for remorse in the lottery's organisers? Somehow this twist is less tragic: it only undermines the randomness of a lottery, not the moral value of keeping families together or of saving women from the deep. Of course any extra death is tragic, but then so is the entire situation: there is no good outcome. But some outcomes are worse than others not for the dead, but for the living. The ill-luck of a spouse choosing to die against the lottery's impersonal "will" is less damaging to the organisers' moral character than the sisters' death was to Holmes and his fellows, because the crew presumably held the sanctity of female life dear in a deep way which no one could feel about the randomness of lotteries' outcomes. The unlucky outcome of a lottery would have been a project failure; in legal terms--and, I argue, in ethical ones--the crewmen's failure in Holmes was the personal sort, and the grounds for an individual criminal prosecution.

There is another reason for preferring random procedures in cases of scarce medical resources. In such cases, or at least when life and death are at issue, the judgement in a sense surpasses our powers of reckoning. As Pascal's Wager argues in the case of eternal damnation, probabilities are irrelevant, dwarfed by the infinity of the possible loss. No odds are telling. Randomisation recognises this explicitly for the parallel modern loss of dying "before one's time." As Paul Freund has written of social criteria for allocating scarce medical resources:

The more nearly total is the estimate to be made of an individual, and the more nearly the consequence determines life and death, the more unfit the judgment becomes for human reckoning...Randomness as a moral principle deserves serious study.²¹

As we will be doing in chapter eight on deterrence, what we are dealing with in conflicts of life with life is a kind of infinite loss. Perhaps it would be more accurate to say that the loss is the equivalent of an infinite one for the loser in the case of scarce resources, whereas in the case of nuclear holocaust it is the equivalent of an infinite one for the whole world, at least potentially. A strict consequentialist would view even this modified claim as risky, since setting an infinite value on human life would require the agent to go on saving lives obsessively, no matter what the damage to her own health or strictures on her other activities. And to a mathematician or a true believer, death is not an infinite loss, perhaps. Nevertheless, it is the nearest we in a secular age come to the unending loss of eternal damnation, and that makes Pascal's Wager considerations relevant.

[A]s eternity [has] receded in the popular consciousness, longevity [has] moved to the fore.

Among medieval Christians, eternity was a socially recognized need; and every effort was made to see that it was widely and equally distributed, that every Christian had an equal chance at salvation and eternal life: hence a church in every parish, regular services, catechism for the young, compulsory communion, and so on. Among modern citizens, longevity is a socially recognized need; and increasingly every effort is made to see that it is widely and equally distributed, that every citizen has an equal chance at a long and healthy life: hence doctors and hospitals in every district, regular check-ups, health education for the young, compulsory vaccination, and so on.²²

In fact, some consequentialists agree that the fallibility of the utilitarian calculus in life-and-death matters makes random selection the fairest procedure.²³ So might deontologists: paradoxically, respect for persons can best be met by this impersonal procedure. There is nothing hurtful about being rejected by a random-number generator, everything painful about being condemned as surplus to society's requirements and standards. Lotteries are not rational, but they are profoundly ethical. If this is accepted, it should give some pause to those who, like Brandt, claim that the prudential or rational and the ethical are coterminous. (It would of course be consistent with Parfit's unwillingness to equate the irrational with the immoral.)

But is this too enthusiastic an acceptance of the lottery principle? Are there no cases in which we would refuse to intervene rather than allow randomness its sway? At least two sets of hard cases can be identified: those involving the very old, and those featuring dependants. I want to argue that the lottery principle should apply in the first case but not in the second. In this I draw

on the work of Jonathan Glover and John Harris,²⁴ but disagree to some extent with both authors' conclusions.

If there is only one dialysis machine or kidney available, and a choice must be made between giving it to an eighty-year-old or a twenty-year-old, most people find the answer obvious enough. Harris calls this common-sense preference for saving the younger patient a form of ageism and argues that it contradicts respect for persons. The eighty-year-old's attachment to life may be as great as the twenty-year-old's. More generally, patients with short life expectancies may be as worthy of receiving the scarce medical resource as those with longer expectancies. This appears at first to be a similar point to Taurek's about the non-additivity of suffering (see chapter three). But Harris is actually willing to give greater weight to the desire for life in the patient with a shorter expectation of it, in this example of a terminally ill person:

Suppose I am told today that I have terminal cancer with only approximately six months or so to live, but I want to live until I die, or at least until I decide that life is no longer worth living. Suppose I then am involved in an accident and because my condition is known to my potential rescuers and there are not enough resources to treat all who could be immediately saved I am marked among those who will not be helped. I am then the victim of a double tragedy and a double injustice. I am stricken first by cancer and the knowledge that I have only a short time to live and I'm then stricken again when I'm told that because of my first tragedy a second and more immediate one is to be visited upon me. Because I have once been unlucky I'm now no longer worth saving.²⁵

I find this argument from ill luck quite compelling; it is disappointing, however, that Harris backtracks from it almost

immediately by introducing an opposing argument, grounded in "fair innings." Whereas the "anti-ageism" argument took the agent's subjective valuation of life as the criterion, the fair innings perspective sets an objective standard, that of the statistically average lifespan. (Harris should properly divide this into male and female lifespans if he wants to avoid disadvantaging women, but this is not my main criticism, although it remains a surprising omission.) Beyond this limit, set at seventy, the anti-ageism argument is not to trespass. That is, no one over that age is to be allowed the scarce medical resource in preference to someone under that age, no matter how intense the older person's desire for life. But the whole force of the argument that we should not necessarily prefer longer expectations of life in allocating scarce resources rested on its bold and probably true assertion that the desire to live does not diminish with age--still less vanish magically at seventy.

Admittedly, desire to live is not the only factor which we have to consider. There is also the good of living beyond twenty, and it might be argued that the seventy-year-old has already enjoyed that. But there is also the good of living beyond seventy, seventy-one, seventy-two, seventy-three and so forth. If life is good, it does not necessarily become any less sweet after seventy merely because seventy is a norm, assuming the patient is not in chronic pain or other distress. On this argument, we ought always to give preference to the younger person, even to a seventy-year-old over a seventy-one-year-old: but Harris does not want to say that.

Actually Harris does not assert that the will to live does fade away after seventy; but he substitutes the objective statistical criterion for the personal, subjective one once the magic age is past. Of course we need some way of deciding between the twenty-year-old and the seventy-year-old which is not merely subjective, since the strengths of their desire to live may be equal. This is exactly the point of my argument in favour of random allocation. But the fair-innings principle offered by Harris seems a piece of sleight-of-hand. The rationale of the fair-innings argument is that "there is some span of years that we consider a reasonable life, a fair innings."²⁶ But who is this "we"? The anti-ageism argument depended on no such animal, and this was indeed part of its attraction: it made no pluralist assumptions about underlying consensus in society, but frankly posited the most basic sort of conflict--opposing and irreconcilable desires to live. This was a valuable contribution in a field where consensus in society has too often been assumed.²⁷

Because people in modern western societies do normally live to a statistical average of seventy says nothing at all about whether they should live to seventy. To argue otherwise is a form of the naturalistic fallacy. Furthermore, statistical independence ensures that no individual has a definite expectation of living to seventy, let alone a right to it. For all these reasons I find Harris's retrenchment incorporating the fair-innings argument less convincing than his original bold assertion that intervening so as to give preference to the young above the old is an invasion of respect for persons. A random principle should apply even here, I would argue--

either a lottery or, more practically but less rigorously, first-come-first-served. To put the matter in Rawlsian terms, youth is not a form of desert any more than is intelligence or class.

Where I am willing to depart from random principles of distribution for scarce medical resources is the case in which duties rather than desert hold sway--that of dependants. In this I agree with Glover rather than with Harris's critique of Glover's position, although my reasoning is somewhat different from Glover's. In the hard case of a choice between someone with no family and the mother of several young children, Glover holds that:

Refusal to depart from random choice when knowledge about their dependants is available is to place no value on avoiding the additional misery caused to the children if the mother is not the one saved...A large part of the case against interventionism is the undesirability of creating a two-tier community, saying that we value some people more than others. This seems obviously objectionable if our preference is based on the belief that one of the people is nicer, more intelligent or morally superior to the other person. But the objection loses a lot of its force when the preference is justified by citing the interests of dependants rather than the merits of the person selected.²⁸

Harris feels that Glover's preference for saving those with dependants does in fact set up a two-tier system, favouring parents--"a systematic preference of those with families over those without...[T]his looks very much like a covert grading of people into the 'haves' and the 'have-nots'--those who have dependants and those who don't."²⁹ Harris even goes on to suggest that a policy of favouring parents of young children might lead people to "acquire" children in order to benefit from "a relatively cheap form of insurance against a low-priority rating in the rescue stakes."³⁰

Clearly Harris views children as a "good" which some privileged souls possess and other less advantaged people do not enjoy. This is highly suspect: having children is a source of economic disadvantage, particularly for women. Mothers do not find the economic cost of having children cheap--in terms of lost earnings during childbearing and childraising years, diminished pension and seniority rights, and the likelihood of returning to work at a lower level or in a part-time position. In 1982 almost half the families with dependent children on supplementary benefit were headed by single parents--the overwhelming majority women. Half of all single mothers were claiming benefit.³¹ It is probably also fair to say that mothers think of children in terms of duties rather than of privileges. This interpretation is not merely personal, but consistent with the fact that women have been found in psychological studies generally to think more in terms of responsibilities than of rights or entitlements.³²

In the hard case under discussion, we should save the mother because she has duties to perform, not because she has any entitlement or right as a mother which might "privilege" her. Harris might argue that the short-term effect is the same, and of course this is true: but the grounds for preferring the mother do not reinforce privilege in the long term, since, I argue, mothers are if anything a disadvantaged group. Fathers are a more difficult issue, since they might claim to be equally bound by duty--the duty to provide. However, this particular responsibility has been the source of privilege, through the trade union policy of the family wage, for

example, which has systematically rewarded male workers and preserved income differentials between men and women.³³ In order to avoid the accumulation of further advantage--for the good reason cited by Glover and accepted by Harris, the avoidance of a two-tier system--I would view fathers of dependent children as having no claim to special treatment unless they do at least half of the actual child care. A 1984 Department of Employment survey reveals that this would encompass at most about half of fathers with children under 16 if only child care is taken into account, about a quarter if equal shares in housework is also required.³⁴

The exception to randomness for mothers and some fathers of dependent children does not threaten the earlier arguments I made for first-come-first served or lottery procedures. No judgement is made of these parents' social worth or of their life expectancies, likelihood of recovery, or other statistical factors whose accuracy becomes increasingly crucial and decreasingly reliable in life-and-death matters. Further, it is hard to see how medical staff could feel remorse at deciding in favour of the parent. Harris might argue that this is merely conventional prejudice, "parentism," another instance of more favourable treatment for parents, but the economic indicators about maternal poverty have already cast doubt on his claim that all parents are economically advantaged through possessing children. Statistics such as those compiled over the years by the Child Poverty Action Group consistently show that having children is the greatest cause of poverty. Therefore no ill luck in terms of unintentionally favouring already advantaged groups can result from

preferring to save primary carers for dependent children--and perhaps by extension carers for the elderly, who now outnumber mothers of dependent children among women. (There again the responsibilities of caring typically result in lowered economic expectations, through giving up work and losing out on pension rights.)

Having demonstrated, I hope, that allocation of scarce medical goods is indeed an ethical rather than a prudential or clinical judgment, I have argued that random allocation--except in the case of primary carers for dependants--is least likely to occasion remorse in health care agents. It may well occasion regret, however, particularly if a ninety-year-old wins out over a twenty-year-old. This might well be a deeply sad occurrence, but not injustice, I would argue--and we need to keep the two straight.³⁵ Nor would it be an instance of moral luck. Provided that moral agents make the decision about allocation according to the right principle--and I argue that randomisation is generally the right principle--they can satisfy Nagel's maxim, I think, however things actually turn out. That is, they have no reason to modify what decision they would have taken in the light of their knowledge of subsequent events.

The decision to randomise treatment draws chance's fire by incorporating its operations. It does not, of course, excuse a deliberate decision to reduce the total pool of resources available. That this has occurred in Britain in the last decade is well documented³⁶. Does moral luck apply at the level of public policy, too? Should a health minister feel remorse and be deemed to have

suffered personal rather than project failure if the media publicise his decision to cut resources and so "allow kidney patients to die?"

At an intuitive level, I feel that the answer must be yes. But a proper analysis of the question would require a deeper consideration of public versus private morality than this chapter can give. In the following chapter I hope to partially rectify that omission.

FOOTNOTES

1. This approach is exemplified in chapter 15, "Justice and the Allocation of Medical Resources," of Raanan Gillon's Philosophical Medical Ethics (Chichester: John Wiley and Sons, 1985), pp. 93-99.

2. Guido Calabresi and Philip Bobbitt, Tragic Choices (New York: W.W. Norton, 1978).

3. Williams, "Moral Luck," p. 25.

4. A sophist might want to argue that the health professional has no duty of care to the patient who didn't get the kidney which was "wasted" on the patient whose good prognosis proved to be ill-calibrated. The health professional has not cared for that second patient and never will--the kidney having died with the first patient. I think this would be an excessively cynical argument, and not one accepted by the majority of nurses to whom I have taught seminars in health care ethics.

5. Ibid., p. 31.

6. Ibid., p. 25.
7. Calabresi and Bobbitt, Tragic Choices, p.19, distinguishing between first- and second-order determinations.
8. Helga Kuhse, The Sanctity-of-Life Doctrine in Medicine: A Critique (Oxford: Clarendon Press, 1987); Harris, Violence and Responsibility and The Value of Life (London: Routledge and Kegan Paul, 1985).
9. Calabresi and Bobbitt describe the criteria which prevailed in the 1960's and 1970's; my impression is that these efficiency-minded, utilitarian standards are seen as less defensible by British health professionals in the 1980's. Certainly the efficiency model is less publicly acceptable these days, as witness the furore over the discharge from dialysis by the Churchill Hospital in Oxford of a vagrant kidney patient on the grounds that he was insufficiently compos mentis to follow the diet and other requirements for successful treatment.
10. Calabresi and Bobbitt, Tragic Choices, p. 184.
11. Ibid., p. 182.
12. Ibid., fn. 90 to p. 182, describing interviews by Calabresi with two doctors who ran dialysis centres in Pisa and Florence.
13. The brouhaha over the Churchill Hospital's decision to discharge a vagrant from kidney dialysis (see footnote 9) may reflect a growing British discomfort over the class implications of a medical model.
14. Calabresi and Bobbitt, Tragic Choices, fn. 93 to p. 183.
15. Ibid., p. 189.
16. Mason Willich, summarised in James F. Childress, "Who Shall Live When Not All Can Live?" in Samuel Gorovitz et. al., Moral Problems in Medicine (second edition) (Englewood Cliffs, New Jersey: Prentice-

Hall, 1983).

17. Calabresi and Bobbitt, Tragic Choices, fn. 111 to p. 188.

18. David Sanders and Jesse Dukeminier Jr., "Medical Advance and Legal Lag: Hemodialysis and Kidney Transplantation," UCLA Law Review 15:367 (1968), p. 378.

19. Most criticism focused on the class bias of the committee's recommendations. That slant is unmistakable, of course, but criteria favouring high earners will also work against women. The committee was willing to give preference to a non-earning housewife with small children, but once these hostages to fortune were grown, older women would have had to take their chances--rigged chances.

20. U.S. v. Holmes (26 Fed. Cas. 360), cited in Childress, op. cit., p. 642.

21. Paul A. Freund, "Introduction," Daedalus, spring 1969, xiii.

22. Michael Walzer, Spheres of Justice (Oxford: Basil Blackwell, 1985), p. 108. In the eleventh year of the Thatcher government Walzer's final sentence rings more than a little hollow, but that does not undermine his parallel between eternal salvation and longevity.

23. See, for example, Glover, Causing Death and Saving Lives, pp. 218 and 223.

24. Glover, Causing Death, pp. 217-227; Harris, Value of Life, pp. 87-110.

25. Harris, Value of Life, p. 89.

26. Ibid., p. 91, emphasis added. On age as a criterion, see also Margaret P. Battin, "Age Rationing and the Just Distribution of Health Care: Is There a Duty to Die?", Ethics v. 97, n. 2, January

1987.

27. This is a frequent failing in Calabresi's writing, for example.

An example is this discussion of why "society"--whatever and whoever that may be--is willing to devote huge resources to rescuing one trapped miner, but not to improving the overall expectation of life by reducing the likelihood of future accidents in other areas, even if the latter could be done more cheaply. "We are committed to 'humanism', to the dignity of the individual, and to human life...Accident law indicates that our commitment to human life is not, in fact, so great as we say it is; that our commitment to life-destroying material progress and comfort is greater. But this fact merely accentuates our need to make a bow in the direction of our commitment to the sanctity of human life (wherever we can do so at a reasonable total cost). It also accentuates our need to reject any societal decisions that too blatantly contradict this commitment. Like free will, it may be less important that this commitment be total than that we believe it to be there. Perhaps it is for these reasons that we save the man trapped in the coalmine." ("Reflections on Medical Experimentation in Humans", in Paul A. Freund (ed.), Experimentation on Human Subjects (London, 1972), quoted in Glover, Causing Death, footnote 2, p. 210. This pluralist liberality with "we" and "our" also permeates the otherwise excellent Tragic Choices.

28. Glover, Causing Death, p. 222.

29. Harris, Value of Life, p. 105.

30. Ibid., p. 106.

31. Anna Coote and Beatrix Campbell, Sweet Freedom: The Struggle for Women's Liberation (Oxford: Basil Blackwell, 1987, 2nd edition), p.

89.

32. Carol Gilligan, In a Different Voice: Psychological Theory and Women's Development (Cambridge, Massachusetts: Harvard University Press, 1982), p. 105.

33. Coote and Campbell, Sweet Freedom, pp. 63-67, 164, 261-262.

34. Department of Employment, OPCS, Women and Employment (HMSO, 1984). There is disagreement between husbands and wives over how much work husbands do--which will hardly seem surprising. However, the women sometimes assess the husbands' contribution more generously than the men do, and that is unexpected.

35. Harris often conflates misfortune and injustice, I feel, as in his claim that it is somehow unjust not to be allowed to live to the statistical norm of 70 (at p. 93).

36. Estimates of deaths through cuts in kidney unit budgets in the early 1980's have been put at 2000 per year (Guardian, 18 October 1983, leader article).

Chapter Six

Secrecy and Moral Luck

The Watergate scandal eventually forced President Richard Nixon from office after his illicit authorisation of surveillance against his Democratic opponents came to light. "Dirty tricks" perpetrated by "a national security culture, protected...by the shield of secrecy"¹ were revealed in The Pentagon Papers; they helped to defeat the Pentagon's ambitions by spreading opposition to the Vietnam War. The British Special Branch was made to look ridiculous after the public learned that it had carried out surveillance against an alternative form of Brownies and Cubs, the Woodcraft Folk. The Thatcher government's campaign to prevent distribution in Great Britain of Spycatcher vastly increased sales of the book worldwide and publicised the claim of the author, former MI5 deputy head Peter Wright, that the security services had actually plotted to overthrow the Wilson government. Mrs Madeleine Haigh discredited the West Midlands constabulary by forcing its chief constable to admit that she had been spied on after her local paper printed a letter she had written criticising Cruise missile deployment. The BBC's submission to undercover government pressures against journalist Duncan Campbell's "Secret Society" exposes helped to prove exactly the point which Campbell had wanted to make. The cancellation of the BBC "Real Lives" documentary involving a Sinn Fein member, under Home Office instructions, resulted in far higher viewing figures for the programme when it was finally shown months later.

In all these cases, secrecy left those in authority with egg on their faces. In the first two examples, those in authority actually lost power, in large part through misjudgements about secrecy. Yet secrecy is generally perceived, particularly by political Realists and some other consequentialists, as a necessary tool of authority. This is an illustrative, subsidiary paradox to the general one about moral luck.

The use of secrecy is normally justified in terms of producing more favourable outcomes: enhanced national security, or room for manoeuvre in negotiations, for example. Yet in all the cases listed above, secrecy resulted in disasters of varying degrees of gravity. Of course, the reply may be that decision-makers could not have foreseen the unfavourable outcomes, or that even if they had foreseen them, they might not have been able to control them. This is exactly where moral luck comes into it. Normally, in theory, at least, we do dislike holding agents responsible for things beyond their control; but at the same time, in practice, we want people to be responsible to some extent for how things actually turn out. In the case of state secrecy there is an additional wrinkle: we particularly want political leaders to be responsible for outcomes, to be accountable, in a democracy. That we do make this demand is clearly something more than an inconsistency in common sense: it is the fulcrum of democratic theory as well as of democratic practice. Different models of democracy accord accountability varying degrees of weight: from intense scrutiny of almost every action of the leadership in the classical Athenian model, to a restricted right of popular review only once every fifteen years in Hayek.² Nevertheless, some form of accountability, however

minimal, is a sine qua non of every model of democracy. But how can leaders be held accountable if their actions are shrouded in secrecy?

It looks as if the use of secrecy is actually an attempt to evade the operations of moral luck, to assert that leaders are not responsible--because they will never be found out--if things turn out badly. But this cannot be a genuine escape from the paradox, merely a sham. There are subsidiary questions here, too, having to do with prudence and rationality. If the outcome is unfavourable--as it was in these cases, by the decision-makers' own standards--was the decision in favour of concealment merely imprudent? Or is secrecy actually morally wrong if it proves counter-productive?

The puzzles about secrecy are increased by an epistemological Catch-22, neatly expressed by a Whitehall source interviewed about surveillance: "The trouble is that you don't know whether a thing is worth looking at until you have looked at it, and then you don't know whether it's going to change in the future, so you have to carry on looking at it."³ If this is true, it raises particular problems for a theory of ethics as information-centred as consequentialism. The information obtained through secrecy is, in this opinion, valueless as a predictive guide to how much more secrecy will be required, and rational calculation will be impossible.

Within consequentialism there are grounds for and against secrecy, even though the traditional raison d'être of state secrecy is deeply consequentialist. I shall discuss these varying consequentialist

perspectives in more detail. But I shall also argue that responsibility cannot hinge on good or bad political fortune, on outcomes alone. Democratic accountability requires that we retain a notion of political agents' responsibility, though perhaps we will have to limit what they are responsible for.⁴ In this case it would seem that Kantianism is more likely to provide a satisfactory position on state secrecy. However Kantianism is rather surprisingly lax on this score, although Rawlsian contractualism is more satisfactory. I shall organise this chapter around these conflicting perspectives on secrecy and moral luck, particularly in politics--although medical confidentiality and truth-telling will also enter the discussion.

First, however, I need to clear away a possible preliminary objection. It may be averred that no ethical perspective could possibly have anything to say about state secrecy because states are not bound by ethical considerations. National security, the usual justification for both internal and external surveillance, constitutes the prior base on which all theories of ethics rest, according to this raison d'état line. Without the state's protection, there would be an internal war of all against all, and philosophers would not survive to tell their Kantian, consequentialist, or Rawlsian tales. Against others of their ilk, states are entitled to take any action they deem necessary to survive.⁵ Now there is a serious question about what actions states are allowed to take against their own subjects, even--or perhaps particularly--in this Hobbesian doctrine. Internal surveillance is less justifiable than external: the sovereign cannot claim the need for self-preservation against his own subjects, since

by definition they have already submitted to him. Yet many sources--including some former secret service officials interviewed in 1988 for the BBC Radio Four programme "My Country Right or Wrong"--assert that the government is more concerned to keep its operations secret from its own citizens than from the outside world, more worried about avoiding accountability than reinforcing national security.

Even more fundamentally, there may be grounds for doubting the empirical accuracy of the Realist claim that subjects' security is best protected by an amoral state. And predictive inaccuracy will be particularly damaging to the raison d'état doctrine, for reasons which I hope will become clear in the next paragraph. But what might be the reasons for scepticism? I have listed a plethora of cases in the first paragraph--instances in which the state got it wrong. More broadly, what if the national security apparatus is a greater threat to our well-being, or even our survival, than we would face in a state of nature? This Orwellian niggle is an ironically realistic doubt for the relatives of the disparicidos in Argentina. What if the British defence and security establishment is more loyal to its American counterpart than to elected UK officials? What is the effect on sovereignty and security of the UKUSA intelligence community, a world bureaucracy with a quarter of a million full-time staff and funds of \$16-\$18 billion per annum?⁶ This security apparatus has been said to rely even more extensively on secrecy than do national defence establishments. Yet its existence may impose risks on Britain, because a hostile power would need to knock out US spying bases in this country even if Britain were not a formal combatant.

As Stuart Hampshire has pointed out,⁷ whether one affirms or denies raison d'état is itself open to moral discussion--and, I would add, to empirical scepticism about whether it produces the outcome which is meant to justify it. Indeed, a principal argument against the doctrine of raison d'état, in my view, is that it is fatally prone to moral luck disease. The criterion for a right action in politics is frankly success, according to Machiavelli. Historians rather than philosophers are allocated the task of making moral judgments, since it is "a historian's yardstick: continuing power, prosperity, high national spirit, a long-lasting dominance"⁸ which measures the ruler's actions. But this leads to insuperable problems about hindsight, making the supposedly pragmatic guide of Realism a profoundly useless guide to political action. The Watergate revelations led to instability, regime change, national self-doubt, constitutional upheaval, and overall governmental weakness. Bugging is clearly wrong, it seems, because this particular "crime" most certainly did not pay. But the Wright memoirs, indicating that MI5 machinations against Harold Wilson were at the least inadequately investigated by the successive Conservative government, may have had no such damaging political effects. Bugging is obviously all right, then--or at least the suppression of information that it took place is morally justified. We are back to the possibility embraced willingly by Williams, that only success could have justified Gauguin's action. That raised problems of hindsight which I called unsurmountable there. They are all the more so when the political future of nations rather than the artistic future of individuals is at stake.

In a few cases, secrecy looks correct at first, but it does not need consequentialism to justify it. This turns out to be related to the absence of outcome luck as a determining factor. In 1973--shortly before Watergate broke--Nixon defended the secret negotiations which resulted in American prisoners' release in these consequentialist terms: "Had we not had secrecy, had we not had secret negotiations with the North Vietnamese...and had we not had that kind of security and that kind of secrecy for this kind of exchange that is essential, you men would still be in Hanoi rather than in Washington today."⁹ (This standard Realist argument is also interesting for the lie which it gives to Realists who say that we only hear about the secret operations which fail. But for every Vietnam prisoner exchange there is a Bay of Pigs: the objections about moral luck cannot be overridden so easily.) In the case of secret dealings to free prisoners, moral luck may indeed not apply--because there is no prior ethical question of right and wrong, no ethical imperative being overridden in the name of national security. No one debates whether it is praiseworthy or blameful to bring prisoners home. If the negotiations had failed, nothing would have been lost in the ethical sense. Therefore if there is an argument against secrecy in this case, it will not have to do with outcome luck. And in the worst case which could have occurred with the use of secrecy, the men would have remained prisoners--the same outcome as in the worst case with full disclosure. It appears that in such cases secrecy cannot actually worsen the outcome--unless it tempts negotiators to take extra personal risks, as it seems to have done in the harrowing case of Terry Waite.

To conclude this preliminary discussion of public versus private morality as it relates to secrecy and moral luck, I should make it clear that there is a third option. Politics may be potentially judgeable by ethical standards; it may be separate from ethics and subject to no standards other than those of success; or politicians may actually be indictable for actions which would not be immoral if committed by private citizens. This last possibility stands raison d'état on its head, whilst accepting the Realist separation between ethics and politics. And intriguingly, it suggests that one reason why politicians should eschew secrecy is that they should be able to foresee that it is likely to lead to abuse. Even more broadly, politicians should beware of policies which are likely to require secrecy--the foremost being nuclear deterrence (see also chapter eight).

Dennis Thompson, for example, asserts that democratic leaders have a responsibility to avoid in the first place policies that are likely to dirty their hands and require concealment. Nuclear deterrence is particularly likely to lead to what Thompson calls "double dirty hands", because it creates an institutional, secret machinery of massive proportions, and because it may engender more risks than it mitigates.¹⁰ Thompson offers the most institutionally-minded and outcome-orientated version of what I identify as the third option, but he echoes Acton's rejection of a Realist's argument:

I cannot accept your canon that we are to judge Pope and King unlike other men with a favourable presumption that they did no wrong. If there is any presumption it is the other way against holders of power, increasing as the power increases.¹¹

Jefferson likewise held that governmental secrecy was least justifiable in

matters of national security. Citizens had a right to full information about any risk of war: "It is their sweat which is to earn all the expenses of the war, and their blood which is to flow in expiation of the causes of it."¹²

I have spent considerable time clearing away what was meant to be a preliminary objection--against equating responsibilities in ethics and politics--but hope in the process to have already illustrated some difficult points about secrecy and moral luck, such as the one implied from Thompson. One other distinction also needs to be dealt with. There are two kinds of arguments about secrecy. One, raison d'état, I have called consequentialist, and so it is for states; but it is only consequentialist for citizens insofar as they accept that the state must employ secrecy and other "dirty tricks" in order to assure its, and their, survival. In other respects raison d'état assumes a deontological morality of absolutes in the private realm. Individuals must be barred from breaking the rules, if there is to be a contrast between their behaviour and that of a sovereign who is not so hampered. The second type of argument about secrecy maintains that the moral enterprise is consequentialist across the board. Both individuals and states should act with a view to outcomes, and there is no separate container in which individual morality is kept spic and span. This view would admit problems of medical ethics such as confidentiality to be weighed up with the same consequentialist scales as public policy matters.

Sidgwick, for example, claimed that secrecy could actually render a wrong action right--in the public or the private realm. "It may conceivably

be right to do, if it can be done with comparative secrecy, what it would be wrong to do in the face of the world."¹³ On this ground, what looks wrong about the British Medical Association's 1986 resolution urging doctors to screen patients for HIV status at random and without obtaining informed consent (see chapter 7) is not that this practice constitutes an assault--which it certainly does, even under the feeble UK law on consent--but that someone in the BMA leaked the recommendation to the press. This peculiar interpretation seems to be reinforced by Sidgwick's characterisation of the correct utilitarian policy: "that the opinion that secrecy may render an action right which would not otherwise be so should itself be kept comparatively secret." This is the precise contrary of the second-order principle advocated by Sissela Bok: that the principles governing secrecy should themselves always be public.¹⁴

Sidgwick's "Government House utilitarianism," as Bernard Williams typifies it,¹⁵ sits uncomfortably with his attempt to categorise truth-telling as one of the values which utilitarianism advocates as producing the best overall state of affairs. (Justice, spontaneous affection, loyalty to friends, and parental caring also fall into this category.) What Sidgwick is grappling with, however, is the paradox of whether utility is best maximised by utilitarians keeping their identity mum. Sidgwick's intention seems to have been to restrict carte blanche in secrecy to "enlightened Utilitarians",¹⁶ to whose ranks he would perhaps not have welcomed Richard Nixon or the BMA in the midst of an AIDS moral panic. What should be kept secret, to Sidgwick, is not politicians' dirty tricks, but the principles of utilitarianism themselves, in rather Masonic fashion. Utilitarianism is

best served when such values as truth-telling become automatic and universal, but they cannot become widely adopted by those who lack the necessary calculating sophistication unless they are justified in old-fashioned non-utilitarian terms, as virtues. This utilitarian modification of the Platonic noble lie is required because utilitarianism may actually result in an overall lowering of moral standards if promulgated with full frankness to the intellectually unfit. The full quotation from Sidgwick is illustrative: "Thus, on Utilitarian principles, it may be right to do and privately recommend, under certain circumstances, what it would not be right to advocate openly; it may be right to teach openly to one set of persons what it would be wrong to teach to others; it may conceivably be right to do, if it can be done with comparative secrecy, what it would be wrong to do in the face of the world."

Sidgwick's position is unappealing, but far less extensive than the raison d'état assertion that state secrecy in the name of national security is always justifiable. Similarly, a consequentialist line on medical confidentiality appears rather harmless--perhaps too harmless. This is one form of secrecy which most of us would favour--provided it were not extended to forbid us to see our own records, which are still broadly regarded as health authority property in Britain, despite some recent Parliamentary changes in the direction of more openness. So far as doctors' obligation to respect confidences goes, the argument is generally consequentialist:¹⁷ better diagnosis and greater trust are likely to result from practitioners' undertaking to keep their patients' secrets. A deontologist would argue that these grounds were not sufficient, however,

and that they had to be underpinned with respect for patient autonomy. This does appear to be essential in practice, since as many as a hundred health professionals and administrators at one American university hospital were granted access to "confidential" records on the consequentialist grounds of improving diagnosis and treatment.¹⁸

The more relativistic consequentialist approach has been unable to prevent erosion of privacy in principle as well as in practice. The General Medical Council's "blue book" spells out eight exceptions to the principle of confidentiality, one of which is the example in which another health professional enters the case, as at the university hospital. Others include the doctor's judgement that a relative or friend should be given access to the information, and further, his view that the "best interests of the patient"--clearly a utilitarian concept--dictate that some third party other than a relative or friend should learn confidential facts. Both these cases would probably be disallowed by a patient-autonomy view, the more strongly because the GMC guidelines require that the doctor should only expose confidential information to outsiders and relatives if the patient has rejected "every reasonable effort to persuade." If the patient so obviously wants to maintain her privacy, a deontologist or rights theorist might argue, that is exactly the case in which her privacy should be most stringently respected.

Again, these consequentialist-inspired exceptions to a consequentialist-based rule of confidentiality allow private records to be opened in order to aid research approved by a "recognised ethics

committee." Particularly because hospital ethics committees themselves operate with very considerable secrecy, this seems very duplicitous. In addition research trials often rely on keeping their purpose secret from patients--and sometimes from on-the-spot researchers, too, in double-blind trials. There are all sorts of scientific and statistical arguments for and against this practice, but no one denies that it overrides informed consent (see chapter 7). One bowel cancer patient with some predicted years to live was entered in such a trial without her or her daughters' knowledge, and died two weeks later; this project had been approved by no fewer than eleven hospital ethics committees.¹⁹

Two other interesting cases are suggested by jury trials and secret investigation of legislative corruption. In each, consequentialism again normally justifies secrecy--though for the sake of what might at first appear a Kantian value, openness. Secret jury deliberations could be defended from within a consequentialist framework, however--as increasing the probability of a fair decision. They could likewise be attacked as reinforcing race bias or the sexist prejudice that rape victims deserve what they get, and indeed I have written elsewhere in an interpretation of the classic study The American Jury that this is exactly what jury secrecy does protect.²⁰

I stated at the beginning of this chapter that there was no single, consistent stand in consequentialism on secrecy, and this observation holds true in respect to consequentialist attitudes towards jury secrecy. If some consequentialists defend jury secrecy on the grounds that it increases the

chances of a fair trial, it is also true that even the minimal amount of secrecy embodied in jury trials has been challenged from a consequentialist perspective when it seems to result in "too many" acquittals. A consequentialist might be willing to balance the extent of jury secrecy against the actual or even the projected rate of acquittals of dangerous criminals. But this would again entail serious problems about foresight and moral luck. How could we know whether this particular case is one in which secrecy should be partial, total, or missing altogether? We cannot know until we know the overall rate of acquittals, but that depends on the outcome in this case. Further, we run into all the difficulties encountered earlier in Nesson's example of the twenty-four guilty prisoners and the one innocent man. The problem here, as there, is that we have to get this decision right this time in an ethical sense, which will not necessarily be the numerical sense. Basic principles of justice are offended by the idea of taking a particular defendant's case as one of a statistical long run, in which we can be content to get the level of secrecy right over time even if not in the particular case.

In the ABSCAM investigation of Congressional corruption,²¹ the FBI used secrecy to uncover secrecy. Undercover agents induced public officials to accept bribes and taped the transactions with hidden recorders. This operation was justified in cost-benefit terms--the level of loss to the nation through white-collar crime, for example, which was estimated at ten times the loss through all street crime. Against these statistics were marshalled arguments that undercover agents only prove the existence of the crimes they generate. But whichever way the cost-benefit equations run in

this case, it does appear paradoxically true that eventual publicity could be had only through initial secrecy.

As in the case of the American prisoners in Vietnam, outcome luck appears not to apply here, not to provide arguments against this particular use of secrecy. The only way in which Operation ABSCAM could be said to have failed would be the case in which a clean-living Congressman rejects a bribe and then discovers that the contact was an FBI agent. But it is hard to see why this would constitute failure for the project or the agent, why any remorse or regret should ensue, or what complaint the Honest Abe would have had against the operation. Indeed, he would probably be chuffed to find his own stock soaring among his constituents once his probity was revealed. In Kantian terms, the Categorical Imperative would prevent his less upright colleagues from complaining that secrecy had been used against them, when they had been covering their own criminal behaviour under the cloak of secrecy.

Now if secrecy is to be defended, it is generally exonerated through consequentialist arguments. This is as much true in the obvious case of national security as in the more esoteric and more difficult examples discussed so far, medical confidentiality, jury anonymity, and investigation of legislative corruption. But consequentialism does not always advocate secrecy. Bentham's "On Publicity" goes so far as to deny that the utilitarian calculus can ever prescribe clandestine actions by states: "Secrecy, being an instrument of conspiracy, ought never to be the system of a regular government."²² "System" may be the operative term here:

rule utilitarianism seems to suggest the most cogent consequentialist arguments against concealment, as a bad overall principle. For example, secrecy is alleged to be addictive, or to produce a brutal attitude in leaders. Habitual truth-telling is a good keep-fit exercise for politicians, Hampshire claims. Indeed, in the Watergate case, Nixon's addiction to secrecy was linked with overall callousness, as revealed in his foul language and bullying on the White House tapes, though whether secrecy produced callousness is unclear.

But again, it is debatable whether utilitarianism even of the rule variety can itself produce any other values. Going further, Hampshire also claims that the utilitarian calculus itself engenders a blunted moral sensibility. I shall not evaluate this large allegation, which entails huge problems of proof, cause and effect. (Even more broadly, Hampshire suggests that "rational" is a term of reproach in ethics and politics, and that "rational computational morality" is a pejorative term.)²³ But the possibility is suggested that we need to move outside consequentialism in order to escape from the moral luck paradox in respect to secrecy. If Hampshire is right, utilitarianism raises objections of brutalisation, to which it correctly directs our attention, but it cannot itself satisfy those objections. We do not want political leaders to become brutalised: we want them to retain a sense of agent-responsibility in some measure. But to the extent that consequentialism shifts the locus of moral value from agent to outcomes, it cannot explain why we should value a responsible attitude in our leaders. It can only explain why we value certain outcomes, or, in

an even more limited fashion, assert that we do value those outcomes and tell us something about how best to achieve them.

Most consequentialist anti-secrecy arguments do focus on concrete outcomes, rather than on attitudes. "The consequences of excessive secrecy are, amongst other things, that it: makes rational debate and discussion of national security issues almost impossible; restricts the effective participation of the public, organised interests, and elected politicians in the policy-making process; violates the principle of democratic accountability; facilitates the concentration of power; undermines confidence in government; and encourages the abuse of power."²⁴ Secrecy in strategic matters is actually said to increase the risk of nuclear catastrophe through fatal misunderstandings. Realists' own criterion of success, as the sole standard of governmental policy is turned against them, with the proposition that security is actually impeded by too little openness.

Greater public information and debate would not actually be harmful to the operations of the security and intelligence agencies. Secrecy has shielded these agencies from full accountability and effective supervision and led to their being less effective and less efficient than they might be.²⁵

But although these claims may well be true, the formulations tend to be a bit bet-hedging. The first quotation refers to "excessive" secrecy but never spells out what that might be. The authors of the second, Jeffrey Richelson and Desmond Ball, go on to state:

In the end the basic dilemma about how to balance the requirements of with those of democratic government cannot be resolved absolutely. Rather, it is a matter of continuous judgment about where to draw a satisfactory line.²⁶

This leaves the field wide open for moral luck (and incidentally for executive tyranny). If a particular use of secrecy succeeds, it will be justified retrospectively, but if it fails, it was wrong to begin with. Of course this leaves no guide for judgement at all, and precious little room for democratic accountability.

Although there are consequentialist arguments both for and against secrecy, in a deep sense it goes against the grain for consequentialism to do anything but condemn secrecy--at least for consequentialism of the modern rather than the Machiavellian sort. There is a tight bond between rationality and most modern utilitarianism, and rational judgement is usually said to demand full information, as in Brandt's first approximation of rationality. (The corollary sometimes drawn from this position is that moral and political disagreement is a function of ignorance. This does not actually follow: conceivably we could all consistently agree from a position of equal ignorance of leaders' motives and acts, but a utilitarian who valued democratic accountability would probably find this insufficient.) Now of course there might be other barriers to voters' possession of full information besides state secrecy; but publicity about leaders' actions and decision processes is a sine qua non for democratic accountability.

Richard Hare, for example, maintains that perfectly prudent preferences rather than actual desires are what should be maximised.²⁷ These are the values which agents would prefer if they were both fully informed and consistent in their thinking. In contrast with utilitarians,

Rawls deliberately builds ignorance into his ethical system, and Williams asserts that we could not rationally desire to be fully informed²⁸-- particularly not about the future, about the time and manner of our forthcoming deaths, for example. But in the extreme utilitarian position taken by John Harsanyi, ethics becomes a branch of the general theory of rational behaviour, along with decision and game theories.²⁹ These methods can of course cope with uncertainty caused by lack of information; Bayesian probability analysis is content to pronounce on samples of one. But typically Bayesians also attach great importance to revising subjective probabilities in the course of a decision on the basis of continually more complete data. Peter Hammond takes it for granted that knowledge should be as complete as the gods will permit. In discussing political decision-making, he notes, "Of course one wants the planner or observer to be as well informed as possible."³⁰ With consistency, he adds that this entails concern when governments deliberately withhold information.

However, utilitarians such as Hammond again have to step over the boundary fences of consequentialism in order to argue for full disclosure. Hammond goes so far as to admit an individual right to know the truth, although he has little time for property rights or rights against the censor. Whilst Hammond believes that utilitarianism can accommodate a right to know, it is hard to believe that Bentham would have thought that any less nonsensical than the other stilt-mounted fictions which he dismissed. In fact, Hammond winds up arguing for an absolute right of access to the truth:

A rather extreme utilitarian might argue that...individuals are no more than the pieces in a utilitarian game, to be manipulated for utilitarian ends, though with their best interests in mind. Leaving

them misinformed is part of that utilitarian game. But it is here where I must at last part company with such extreme utilitarianism, and recognise that individuals certainly have a right to be fully informed, or at least to acquire as much information as they wish to. One can also insist that the social welfare function respect the right of each individual to know as much of the truth as possible, within certain cost limitations.³¹

This rather grandiose formulation gives us little guidance in practice. What would Hammond make of the patient who does not want to know the truth about her condition, although telling her might be required in order to get her consent to a possibly life-prolonging operation? Perhaps her right not to know might flow from the right to "acquire as much information as individuals want to;" but it is hard to see how, in this case, not knowing would be rational or utilitarian even in a non-extreme sense of the latter term. Similarly, what would Hammond make of the conflicting rights of the patient who does not want his family to see his file, and the relatives, who presumably have a "right to know" so that they can make necessary arrangements? Perhaps Hammond would have to say that the patient has no right to own exclusive use of the information; or he might have to fall back on a liberal balancing of individual versus societal rights.

Bernard Williams makes a further distinction foreign to utilitarianism, in clear if curt fashion:

One significant point is that while transparency is a natural associate of liberalism, it falls short of implying rationalism. It is one aspiration, that social and ethical relations should not essentially rest on ignorance and misunderstanding of what they are, and quite another that all the beliefs and principles involved should be explicitly stated. That these are two different things is obvious with personal relations, where to hope that they do not rest on deceit and error is merely decent, but to think that their

basis can be made totally explicit is idiocy.³²

This contains the kernel of a solution to the moral luck problem as embedded in issues about medical secrecy. Again, it looks as if the key is limiting what agents are responsible for--even though Williams himself does not offer this solution to the general paradox. We need to negatively prohibit the withholding of information or the giving of false information by health professionals or authorities--not by individual patients from the authorities or medical staff--but not to enshrine a positive right to know, such as Hammond posits. This limits individual agents' responsibility for how much they are obliged to tell others, retaining some privacy. Thus the patient who wants to keep his condition from his family is not responsible for the likely outcome, that proper arrangements will not have been made on his death. Nor is the medical professional at fault for not forcing information on a patient who "doesn't want to know", although he is in the wrong if he uses this as an excuse for overlooking diffident or fearful requests for more information from patients lacking the requisite assertiveness for dealing with some doctors.

It will be seen that I have somewhat exceeded Williams's own bounds, and shortly it should also become clear that I am going beyond the Kantian requirements in respect to openness. Williams's formulation actually appears to parallel the Kantian distinction between lying and withholding information. In Kant keeping back information does not necessarily vitiate the Categorical Imperative, whereas breaking promises and lying do undermine the essential structure of a moral system. Hence the separate subject matters of Sissela Bok's two books, Lying³³ and Secrets, written

from a broadly Kantian perspective and making the expected distinction in these terms:

The central themes of the two books--lying and secrecy--intertwine and overlap. Lies are part of the arsenal used to guard and to invade secrecy; and secrecy allows lies to go undiscovered and to build up. Lying and secrecy differ, however, in one important respect. Whereas I take lying to be prima facie wrong, with a negative presumption against it from the outset, secrecy need not be. Whereas every lie stands in need of justification, all secrets do not. Secrecy may accompany the most innocent as well as the most lethal acts; it is needed for human survival, yet it enhances every form of abuse.³⁴

Secrecy is then less firmly prohibited than lying, to a Kantian; but secrecy opens up moral luck dilemmas. Does a Kantian argument cope with them adequately?

Whereas a Machiavellian might argue in favour of secrecy on strategic grounds, another sort of consequentialist could equally well argue that honesty is sometimes the best policy. (This latter was the kind of reasoning behind the GMC blue book's exemptions to medical confidentiality: sometimes openness is dictated by the "best interests of the patient," which demand that a third party be given access to confidential information. Richelson and Ball also averred that a consequentialist goal, greater efficiency, would be better served by less governmental secrecy.) But both base their claims on expected utilities or payoffs, and their disagreement is over the probability of success through secrecy and the possible severity of the losses which a clandestine policy might entail.

To Kant, honesty is better than any policy, and this dictum cannot be empirically proved or disproved.³⁵ Although in the personal realm withholding information is less heinous than lying, there is an additional

variable in the political sphere. The state uses secrecy to hide tyranny, and secrecy is in turn an indicator of misgovernment. Perhaps this is less true to a modified Kantian like Bok, who regards the jury system, for example, as a guarantee of individual justice, though justice obtained through a form of secrecy in the public sphere. But Kant rejects secret treaties, and of course the doctrine of raison d'etat--although he retains the Machiavellian distinction between public and private morality insofar as he assigns publicity to be the keystone of one and universalisability of the other. In the appendix to On Perpetual Peace Kant rejects Machiavellian cynicism about the sovereign's power to break treaties, as well as to draw up accords surreptitiously. Promise-keeping and truth-telling are as imperative for states as for individuals, if not more so as far as withholding information goes. Kant is unimpressed by the counter-argument that this could mean political suicide. This outcome-blind pronouncement can be summarised as "Let justice prevail though the world perish for it."³⁶

This begins to sound very much like what Williams calls "moral self-indulgence."³⁷ There is a very real problem about how justice can prevail if the human world has disappeared. Yet oddly, Kant may sometimes be a better strategist in his stickling for openness than Realists who are too doctrinaire in their loyalty to the clandestine. As I discuss at greater length in chapter 8, nuclear deterrence actually requires transparency, though the military establishment draws a veil over such matters in the name of national security. But each side needs to know that the other's Domsday machine works, and will be used. Furthermore, bluff and secrecy,

such as the US employed in the late 1940's--when it really had only thirteen atomic bombs in all--produce serious risk of misjudgement by the enemy, escalation of the arms race, and greater likelihood of war.

Does this necessarily mean that the deontological approach to secrecy is the right one? A modified Kantianism does avoid problems of hindsight and epistemological absurdities of the sort highlighted by the quotation from the Whitehall source at the start of this chapter. But Kant's own writings on state secrecy do, for once, support Hegelian charges of Tartuffesque hypocrisy. With what has been called his "characteristic two-sidedness,"³⁸ Kant glorifies both individual freedom and state power. The law, being a priori and derived from practical reason, must be obeyed, and only the sovereign has the authority to decide the law. It is up to him to act in line with the general will; the subject's duty is to obey the law as it stands. Because disobeying the law is not universalisable, it offends against the Categorical Imperative. But Kant's abstract argument ignores the patent non-universalisability of some actual laws on secrecy. For example, the British Official Secrets Act of 1911 would have led to absurdity if the very broadly worded Section 2 had been obeyed to its extreme. As the Franks Committee remarked wryly in its 1972 report:

We found Section 2 a mess. Its scope is enormously wide...A catch-all provision is saved from absurdity in operation only by the sparing exercise of the Attorney General's discretion to prosecute.³⁹

Further, although Kant conceives of the ideal state as republican, he makes no provision for keeping it that way. He considers separation of powers to be the test of a government's conformity to the original contract, and tyrannies to be typified by amalgamation of the legislature and the

executive. Yet he lacks the awareness possessed in his own time by Madison of the ways in which the executive can emasculate the legislature--often through deliberately withholding or even falsifying information, as in Reagan's dealings with Congress over Iran and Nicaragua, and in the Belgrano affair.

If we are to adopt a Kantian approach to state secrecy in an effort to minimise the effect of moral luck, I do not think we should adopt Kant's own. In Kant, external liberty is equated with freedom from obeying any law to which the agent might not have consented. More precisely, a monarchy may legitimately replace a republic, for example, so long as the general will could conceivably have favoured this course. In respect to official secrecy, the government will be in the right unless I can prove that I could not possibly have consented to the withholding of information. This shifts the burden onto the individual, although it does appear to lessen the problems of hindsight and give chance less sway by making it less important how things actually turn out. If I could not possibly have consented to the use of secrecy, then it was wrong even if it resulted in a favourable outcome.

What we need is a combination of that refusal to let responsibility hinge on outcomes with a stronger line on decision-makers' responsibilities are. Oddly, exactly by incorporating chance and luck, Rawls is able to shift the burden back onto the state, where it belongs. In a further paradox, Rawls uses deliberate ignorance⁴⁰ to investigate allowable secrecy; but the intention is to enhance rationality. In Rawls there is a

powerful meta-theoretical reason for openness. Our deliberations in the Original Position would be pointless if we were barred from knowing how well or badly things actually turned out once the Veil of Ignorance was lifted. It would be deeply wrong if society were so opaque that gender roles, to use an example from Williams,⁴¹ rested on mis- or even dis-information. If women turned out to occupy an inferior position once the Veil was lifted, agents should have chosen this possible outcome rationally from behind the Veil--rationally in the more-minimal-than-Brandt's sense that myths should have played no part. That is, there should be no false information, even if not all true facts are known. This would rule out the conventional stereotypes justifying wage differentials--such as greater male "strength," which is inaccurate in terms of longevity if not of muscle power, and which bears no relation to the higher salaries for (mainly male) managers than for (mainly female) secretaries when both groups are non-manual workers.

Publicity is indeed central to Rawls,⁴² but his requirements fall short of the peremptory right to know suggested by Hammond. In Rawls the link between transparency and democratic accountability is properly clear, but this is no longer a value extrinsic to the theory, as in utilitarianism. It follows from the constraints of the Original Position: if governments are not accountable and open, how can we know whether we got what we decided on behind the Veil? Without publicity, whether we get what we wanted is entirely open to luck, and even more deeply, whether we find out what we got is also open to chance and the arbitrary powers of decision-makers.

Rawls's liberal focus on consent makes a good link to the next area in which I want to explore moral luck, that of informed consent in health care ethics. I also intend to reiterate some of the themes which have dominated this chapter and which will carry forward into the next, particularly rational calculation, predictability, and hindsight. Again in the following chapter, I will look at these issues which flow from the question of moral luck and see how they might be handled by competing moral traditions.

FOOTNOTES

1. Senator Mike Gravel, introduction, The Pentagon Papers (Boston: Beacon Press, 1971), I, p. xi.
2. F.A. Hayek, Economic Freedom and Representative Government (London: Institute of Economic Affairs, 1973). Hayek proposes a legislative assembly which is elected for fifteen years, with one-fifteenth of the membership turning over annually. He also recommends that "At each election the representatives should be chosen by and from only one age-group so that every citizen would vote only once in his life, say in his fortieth year, for a representative chosen from his age-group." (p. 20)
3. Nick Davies and Ian Black, "Security's Legitimate Tasks--and How the Guidelines Are Bent," "The Watchers," Guardian, 19 April 1984.
4. With the waning influence of the Realist school of politics, political theorists are increasingly willing to ascribe moral responsibility in politics. Dennis F. Thompson's Political Ethics and Public Office (Cambridge, Massachusetts: Harvard University Press, 1987), discussed later

in this chapter, is indicative of this trend, especially chapter two, "The Moral Responsibility of Many Hands."

5. A more sophisticated summary of this Machiavellian argument is given by Stuart Hampshire in "Public and Private Morality," chapter two of the volume of the same name edited by Hampshire (Cambridge: Cambridge University Press, 1978), p. 49.

6. Jeffrey T. Richardson and Desmond Ball, The Ties that Bind (London: Allen and Unwin, 1985), p. 311.

7. Hampshire, p. 50.

8. Ibid.

9. Quoted in Sissela Bok, Secrets: Concealment and Revelation (Oxford: Oxford University Press, 1986), p. 181.

10. Thompson, Ethics and Office, p. 33 ff.

11. John Emerich Dalberg, Lord Acton, letter to Mandell Creighton, in Essays on Freedom and Power, cited in Bok, Secrets, p. 298.

12. Jefferson, letter to Elbridge Gerry, 26 January 1799, cited in Bok, Secrets, p. 191.

13. Henry Sidgwick, Methods of Ethics (London: 1874), pp. 489-90.

14. Bok, Secrets, p. xvi.

15. Williams. Ethics and Limits of Philosophy, p. 108.

16. Ibid.

17. Gillon, Philosophical Medical Ethics, p. 108. In France a deontological line is taken, but neither autonomy nor patients' rights is respected: confidentiality can never be overridden, not even by the patient when it would be in her own interest.

18. M. Siegler, "Confidentiality in medicine: a decrepit concept," New

England Journal of Medicine, 1982, 307: 1518-21.

19. Faulder, Whose Body Is It?, p. 9.

20. Donna Dickenson, "Juries, Secrecy and Democracy," chapter 3 of block 6 reading guide, course D308, Democratic Government and Politics (The Open University, 1988).

21. Reported in Bok, Secrets, pp. 266-278.

22. Jeremy Bentham, "On Publicity," Works, II:310-317, quoted in Bok, p. 171.

23. Hampshire, "Public and Private Morality", pp. 23 and 48.

24. Anthony McGrew, "National Security as a challenge to liberal democracy," in paper 14, D308 (1987).

25. Richelson and Ball, Ties that Bind, p. 311.

26. Ibid. To be fair, Richelson and Ball do maintain at least one firm rule: officials of the US and UK governments must reveal as much information deliberately to their citizens as they do inadvertently to Soviet agents! However, citizens do not know how much Soviet agents know. We are still unsure whether the past head of MI5, Sir Roger Hollis, was a Soviet agent.

27. R.M. Hare, "Ethical Theory and Utilitarianism," in Sen and Williams (eds.), Utilitarianism and Beyond, pp. 23-38, at p. 28.

27. Williams, Ethics and Limits of Philosophy, p. 57.

29. Harsanyi, "Morality and the Theory of Rational Behaviour," in Sen and Williams (eds.), Utilitarianism and Beyond, pp. 39-62.

30. Peter J. Hammond, "Utilitarianism, Uncertainty and Information," in Sen and Williams (eds.), Utilitarianism and Beyond, pp. 85-102, at p. 99, emphasis added.

31. Ibid., pp. 101-102, emphasis added.
32. Williams, Ethics and Limits of Philosophy, p. 102.
33. Sissela Bok, Lying: Moral Choice in Public and Private Life (Hassocks, Sussex: Harvester, 1978).
34. Bok, Secrets, p. xv.
35. This useful point is made by Pierre Hassner in his article on Kant in Leo Strauss and Joseph Cropsey (eds.), History of Political Philosophy (Chicago: Rand McNally, 1972, second edition), p. 567.
36. Ibid.
37. In the article "Utilitarianism and Moral Self-Indulgence," in Williams, Moral Luck, pp. 40-53.
38. Hassner, "Kant", p. 561.
39. Report of the Franks Committee, quoted in Dickenson, reading guide to block 6 of Democratic Government and Politics, p. 9. The new act passed in 1989 retains a very broad ambit in the sense of eliminating any defence for publication, even if revelation would be in the national interest, although it limits the scope of the information covered to exclude the classic absurdity covered by the 1911 act: who supplies the cakes in the Ministry of Defence canteen.
40. As Williams points out, however, agents' ignorance is less radical in Rawls than in Kant, where individuals choose almost as disembodied wills and no more. Those in Rawls's Original Position have some knowledge of the social sciences--a little basic psychology--though none of history.
41. Williams, Ethics and Limits of Philosophy, p. 101.
42. See Rawls, A Theory of Justice, p. 133 and section 29.

Chapter 7

Responsibility, Risk, Rationality and Consent

The question of informed consent has been called "the ethical issue in medicine today"¹--less sensational than surrogate motherhood, and less apparently concerned with life and death than kidney allocation, but more likely to affect almost everyone at some point. Consent is a promising instance for illustrating the practical implications of moral luck, together with risk analysis, uncertainty, rationality and associated concerns, to a lesser degree. In broad terms, I have been asking whether "luck" and "moral" are incompatible: whether or not there is a purely ethical area unaffected by the uncertainty of outcomes, whether risk undermines ethics. The first question to ask about informed consent in health care ethics, then, is whether it eliminates questions about risk. If a patient gives her consent to a medical procedure, is the doctor free of moral blame if that procedure turns out badly?²

Clearly this is a question about moral luck, and it will be the main one in this chapter. Certainly in legal terms, failure to obtain the statutory consent proforma may lay the health professional open to a civil charge of assault. But even in legal terms we need to make some further distinctions: between negligence and consent, for example. In general, consent, unlike negligence, concerns a procedure which has been performed correctly but which has ended in an unfavourable outcome.³ For example, one

consent case involved an optional pain-relieving operation with a .01 probability of paralysis and a .02 probability of damage to the nerve roots. The surgeon was accused of not having informed the patient of the apparently lesser risk of paralysis, which did occur. It was alleged that this was a failure to obtain a fully informed consent. However, no one alleged that the operation had been performed incompetently. Through sheer statistical necessity, some outcomes will be unfavourable, and the question raised by informed consent is who bears responsibility for those outcomes.

A simplistic but promising answer to my initial question might be that it is the giving of informed consent which stops the probability machine rolling, which shuts out questions about moral luck and risk for the health professional. If the patient has given her consent, the ethical and legal requirement--if I may momentarily collapse the two--has been fulfilled. No matter how badly the procedure turns out, there is no cause for remorse in the doctor, although of course there will be grounds for regret.

To put it another way, the ethical realm is bounded by the giving of informed consent, and there the prudential realm begins, the area of decisions which we are content to get right in the long run. If the procedure goes badly this particular time, the health professional is still free of ethical or legal blame so long as he has obtained informed consent from the patient beforehand. The particular outcome will be a project rather than a personal failure for him, and moral luck will not have undermined his agency. Nor will he have neglected the duty of care which he owes to each patient--and which, as I noted in chapter 5, forbids him to

write off a failure as a matter of "better luck next time". The duty and the responsibility will have been taken care of through the obtaining of informed consent, although of course the doctor will still be bound by the normal standards of medical negligence in the performance of the operation.

Even John Harris, who is extremely outcome-orientated in his elucidation of negative responsibility,⁴ views informed consent and full disclosure of information to patients as absolute requirements. Breach of the obligation to obtain consent cannot be palliated by the procedure's turning out well; nor is the responsibility weakened in cases which other writers find tricky, such as random clinical trials or fatal conditions whose disclosure might induce patients to commit suicide.⁵ It is a blanket duty required by respect for persons--an oddly Kantian claim for Harris, although he also makes a subsidiary utilitarian assertion that honouring the autonomy of individual patients enhances the well-being of the community.⁶

But here Harris is in line with most writing on consent, which rarely accepts a strict consequentialist line. With a few exceptions, which I shall discuss towards the end of the chapter, U.S. law is particularly eager to avoid any utilitarian balancing in this area, and the language is typically rights-slanted:

Every human being of adult years and sound mind has a right to determine what shall be done with his own body.⁷

Anglo-American law starts with the premise of thorough-going self-determination. It follows that each man is considered to be master of his body, and he may, if he be of sound mind, expressly prohibit the performance of life-saving surgery, or other medical treatment. A doctor might well believe that an operation or form of treatment is desirable or necessary, but the law does not permit him to

substitute his own judgement for that of the patient by any form of artifice or deception.⁸

Respect for the patient's right to self-determination on particular therapy demands a standard set by law rather than one which physicians may or may not impose on themselves.⁹

[E]very patient...shall be provided by the physician [with] the right to informed consent [and]...in the case of a patient suffering from any form of breast cancer, to complete information on all alternative treatments [to mastectomy] which are medically viable.¹⁰

But the great irony here is that generally the law is brought into play only if the procedure turns out badly.¹¹ How else could a patient obtain standing to sue? That the operation or treatment ended disastrously is not sufficient proof of failure to obtain consent, of course, since negligence may be involved. (However, an unhappy outcome is also insufficient proof of negligence: it must also be shown that the negligence caused the injury directly, and that the health professional was operating at a level of competence below the average for his or her profession, though not necessarily at an ideal level.) Nevertheless, outcome ill-luck is a necessary practical requirement for launching a lawsuit.

There is one exceptional example of responsibility for wrongdoing having been claimed against a health professional even though the outcome of the case was favourable. This was a case involving alleged negligence, but its interest for consent cases as well lies in the rarity of prosecuting a health professional when the outcome has turned out well. A district health authority attempted unsuccessfully to discipline a consultant obstetrician and lecturer, Mrs. Wendy Savage, on grounds of professional incompetence in five obstetrical cases. Three of these had actually turned out favourably. When Mrs. Savage argued that these cases

had clearly been dredged up by trawling through files, the opposing barrister retorted, "That is also the answer of the driver who rounds a blind corner on the wrong side of the road." However, the inquiry subsequently found for Mrs. Savage, apparently on the grounds that her average level of competence was actually above that of her colleagues, who were not being disciplined, and that thus negligence was not proven. The issue of outcome luck was not the crux of the finding.

Apart from this very partial exception, both American and English law are generally consequentialist in practice though not in theory, insofar as no action for damages is possible unless the case turns out badly. In this sense the health professional does not need the protection which informed consent offers--by transferring responsibility to the patient--when outcome luck is favourable. But when the outcome is bad, the remedies available in American and English law differ fundamentally.

The American law of consent attempts to exclude moral luck by requiring all risks to be disclosed, very broadly speaking.¹² In contrast English law tries to eliminate the paradoxes of probability and the vagaries of moral luck by not requiring that the patient be informed of any risks. American law concentrates on consumers' rights, English law on providers' duties. Why are the two approaches so contradictory?

American populism versus English elitism may well be part of the answer, but another component may be the tension between "informed" and "consent". I suggested earlier that the patient's consent does at first

appear to make the morality of treatment risk-free for the health professional, to absolve him of blame and liberate him from remorse if the procedure turns out badly. "Consent" seems to get rid of risk. But the consent must be informed, and that means telling the patient about the risks. These two components of informed consent pull different ways, and in a sense the discrepancy between American and English law reflects that difference.

A well-known example of the more rights-centred and patient-orientated approach to consent in American case law is Cobbs v. Grant,¹³ in which the court found that a surgeon, Grant, had failed to disclose a chain of low-probability events to an ulcer patient, Cobbs. The overwhelmingly unlikely simultaneous occurrence of all these events--injuries to the spleen, development of a new gastric ulcer, and premature absorption of sutures--catapulted the patient into hospital three more times and required the removal of his spleen and half his stomach. Cobbs's GP had discussed with him the risks inherent in general anaesthesia, and Dr. Grant had explained the nature of the operation. In court Grant's lawyer argued that it was not normal medical practice to divulge any more than this--the Bolam standard (see footnote 17) to which English courts still adhere. This argument in terms of how doctors define their duties rather than how patients construe their rights was rejected by the court. Noting that "the patient's right of self-decision is the measure of the physician's duty to reveal," the court insisted that the surgeon should have disclosed "all information relevant to a meaningful decision process." This very wide construction of "informed" has engendered all sorts of problems about decision-making and

rationality. Before going on to examine them, particularly in terms of my earlier concern about prudential decision-making and risk, it might be instructive briefly to compare the English approach.

Thirteen years after the Cobbs case, the highest English court upheld the argument which the Californian justices had rejected--that it is entirely up to the doctor to decide how much information his patient requires. If a doctor fails to disclose a serious risk in treatment, the patient's consent is nevertheless valid in English law. In the case of Sidaway v. Board of Governors of the Bethlem Royal Hospital,¹⁴ the holding was even more extreme. A patient who alleged that she would not have given her permission if she had been told all the risks was held to have given informed consent regardless.¹⁵ The High Court, Court of Appeal, and House of Lords all disallowed the patient's claim for damages, which Mr. Justice Skinner in the High Court had assessed at £67,000, for severe disability following an elective operation on her cervical spine to relieve pain. Rather surprisingly, the reason for denying Mrs. Sidaway's claim was not that she might have been a victim of hindsight bias, people's tendency to assert that they had a much more certain knowledge of future outcomes at the time they were making a decision that tests at that time reveal they actually possessed. (For example, a group of subjects were asked to "forecast" the outcome of a historical battle with which they were unfamiliar. Although the average probability assigned to victory for the apparent underdogs was .20, subjects claimed they had assigned fifty-fifty odds when told that the underdogs had won.) Instead, the English courts phrased their finding in terms of risk and responsibility.

Mrs. Sidaway denied that the surgeon--whom she described as a "man of very, very few words"--had revealed any possible side-effects or risks of the operation to her, or indeed that he had informed her that the treatment was elective. Between the operation in 1974 and the judgement of the High Court, the neuro-surgeon had died. In stunningly circular fashion, the court chose to assume that he had followed standard practice--which was, of course, the very issue in question. Normal procedure would have been to inform the patient of the possible risk of damage to the nerve roots ($p=.02$) but not of possible damage to the spinal cord ($p=.01$). It was the latter--with half the likelihood but an infinitely greater loss, partial paralysis--which occurred in this case.

Now there are many arguments against allowing medical personnel to decide what risks they should mention to patients, the usual ones anti-paternalistic.¹⁶ There are also consequentialist grounds for distrusting the "medical standard" approach in Sidaway: it becomes the profession's interest to narrow the scope of "normal" disclosure.¹⁷ But perhaps the most ironic and telling argument is that it is against the doctor's own interest in avoiding remorse and personal failure to wield so much apparent power as the Sidaway case allows. This is not a paternalistic argument against medical paternalism. Nor is it about what the rational course would be for the physician. It is an expressly ethical question, hinging on the transfer of responsibility through informed consent.

After the Court of Appeal had found against Mrs. Sidaway, the editor of the Medico-Legal Journal wrote:

[In the Court's view] [i]t is a matter for the doctor to decide

how much or how little he should tell his patient, taking into account all the circumstances of which he knows, including the patient's true wishes. As Sir John [Donaldson, Master of the Rolls, giving judgment] pointed out, though many patients may say they want to be told all, it is clear that some of them don't. It is for the doctor to divine which is the case...The doctor would seem to have to be a cross between a detective, a fortune teller, and an emotional and rational prop in some cases...The responsibility is heavy.¹⁸

Just so. Doctors might be expected to prefer telling their patients as much as possible in order to absolve themselves of responsibility if the procedure misfires. Yet this rarely seems to be the case--perhaps because greater openness in the American medical profession has coincided with an increase in litigation. But perhaps both phenomena are symptoms of a third, causative one: greater knowledgeability and assertiveness in patients.

The moral mathematics in Sidaway raises deeply problematic issues, and it is understandable, if not condonable, that English courts have tried to skirt them. Should patients be informed of all risks with a probability greater than .01? Why not .02? How about .05? Should it depend on how much time doctors have, and that in turn on how many junior doctors the government chooses to train? How can we know what the probabilities are before the disasters occur, in the case of new treatments?

Mrs. Sidaway was in severe pain before the operation, even though it was an elective one. If the pain is sufficiently great, is it irrational to accept a high level of risk? Even if it is imprudent for the patient to accept a high level of risk, how does that affect the doctor's responsibility?--bearing in mind that here, as in the case of scarce medical resources, it is the health care professional who is the moral agent, not the patient. Borrowing the idea of identity from Bernard

Williams¹⁹--though not the direction of his argument--could we ask whether a person in great pain is almost a different person from the "same" person after the pain-relieving operation? What about the institutionalised and the mentally ill? Can we assume rationality in those cases? Is it patronising to assume the reverse? Does the very fact of hospitalisation lessen the ability to give an "informed" and rational consent, insofar as hospitalisation produces dependency?

Most important of all is the question of what risks it is permissible for the health care professional to impose on the patient, if any. Derek Parfit²⁰ might not accept my contention in the previous paragraph that the ethical questions here concern the doctor's agency, not that of the patient. He would probably say that there is an ethical question about the patient as well: she is not allowed to treat her future self with any less respect than the doctor must show to her. As I mentioned in my discussion of Parfit's smoking example, that would be to collapse the distinction between the ethical and the prudent, and indeed this seems to be Parfit's intention in that instance. In contrast, I began by shrinking the realm of the ethical in informed consent cases to the obtaining of the consent itself, consistently with my view that we can only defend a notion of moral agency against moral luck if we limit what agents are responsible for. (I think we should also limit who is the agent, since this follows from denying that the patient's imprudence is necessarily immoral, as I do.) Even if the treatment turns out to have an unfavourable outcome, medical and nursing personnel have not acted unethically so long as they have obtained consent, I hypothesised. They might or might not have acted

imprudently--not necessarily, if the unlucky outcome was the statistically inevitable one in 100 times a 99% successful procedure goes wrong, and if no better odds than 99-1 are available from other treatments.

Parfit's argument in the smoking example seemed to extend the realm of the ethical almost infinitely in all directions. Is it morally wrong to impose any risk on either oneself or others? Parfit seemed to answer yes, unless, perhaps, there is a countervailing utility. Yet the smoker clearly believes that smoking does bring him countervailing utilities, unless we assume that his behaviour is deliberately self-destructive. Is he merely inept at calculating expected values? Clearly we are back in the grey area of rationality, prudence and ethics. We will need to examine what standards have been suggested for rationality in informed consent--along with the difficulties which the operations of probability create for rational decision-making.

Before I go on to discuss rationality and consent, however, I want to consider other possible replies to the question of what risks health professionals are entitled to impose on patients. As I mentioned briefly in chapter five, some medical staff employ a consequentialist argument that they are under a moral obligation to administer the treatment with the highest probability of cure--that they are not entitled not to impose the risks associated with that method. This claim is least controversial, though still very debatable, when the procedure is presented as the minimum necessary for restoring a mental patient, for example, to rationality and autonomy, so that she can then give an informed consent or refusal to

further treatment.²¹ More dangerously, the line about overriding autonomy so as to confer autonomy has been used to justify vetoing a patient's refusal to consent to ECT, or even psychosurgery. This reasoning may appear Kantian at first, but it actually makes the patient a means towards the end of her own autonomy. (This assumes that the mentally disturbed patient is still a moral being, of course; if one equates moral personality with rationality, the objection disappears, but I am not very happy about doing that.)

Another sort of consequentialist reasoning tries to defend withholding information from the patient or administering a placebo without the patient's knowledge, on outcome-orientated grounds. In one case,²² therapists successfully lessened a patient's drug addiction by substituting a placebo for the drug, without his knowledge. When informed at last, the patient was initially angry, but later came to judge the "treatment" with a placebo a success in his own terms. Because the outcome was fortunate, I would argue, the therapists were able to say: "We felt ethically obliged to use a treatment that had a high probability of success...To withhold the procedure may have protected some standard of openness, but may not have been in his best interests."²³ But what if the procedure had failed, and the patient had discovered the deception? How could the therapists have appeased his justifiable anger unless the placebo had worked? And unless treatment by placebo had a 100 per cent rate of success, how could the therapists have known in advance that it would work?

In this case the therapists were simply withholding a drug, so that the responsibility appears negative--by contrast with another case in which radiography, active treatment, was performed without genuinely informed consent. (I think, however, that the complex issue of positive versus negative responsibility can be avoided: in both cases medical staff had a positive obligation to obtain consent, which they failed to do.) But again, in this second case, the medical professional argued that he was justified by results. This put him in the looking-glass-world position of claiming that he had acted in the patient's best interests when the patient had died. It is a compelling case, however, since to the best of the medical professional's knowledge, the rate of failure for the procedure was zero.

The case involves a hospital radiologist who never warned patients that they might have a fatal reaction during urography because he thought it would do no good. Although he presumably obtained the required consent proforma, demanded by statute, the consent could hardly be called "informed" if such a serious risk remained unexplained. But the death of one of his patients left him unmoved in his conviction that he had acted for the ethical best:

I have done 6000 to 8000 urograms in the past thirteen years and no one has ever had a fatal reaction. We have been doing urograms at this hospital for at least 25 years and no one has ever had a fatal reaction...Because the indications for urography were great and the chances for a reaction were remote I am sure I would have convinced Mrs. E...to have the procedures. She would then have had the reaction and died, and the fact that I warned her would have done Mrs. E...absolutely no good.²⁴

Williams suggests that the attempt to evade moral luck is so central to an ethical way of thinking that we may prefer to abandon the moral

enterprise altogether if we cannot escape from moral luck. The radiographer is doing just that: denying that there is any ethical issue, since moral luck is inescapable here. How can he excuse his failure to obtain informed consent as having produced the best available outcome when the patient has died? How can he be so certain that he would have "convinced" Mrs. E. to have the operation? (Indeed, his use of the term "convinced" says little for his respect for patients' autonomy.) Perhaps when apprised of the gladsome tidings that there had been no deaths during urography for 25 years in this hospital, Mrs. E.--if she suffered from Gambler's Fallacy as well as her clinical condition--would have remarked that it was time for a death, then. Would her response necessarily have been irrational? Even if Mrs. E. had not died, would the radiologist have obtained the best outcome by treating her as a child? Not if autonomy is valued in itself: in that case, the best outcome needs to be considered holistically, as including the right way of making the decision as well as the clinically desirable outcome. The aim, in the cogent view of one nurse, is not to spare patients fear--which is impossible--but to allow them to "worry intelligently."²⁵

It is now time to discuss rationality and consent in greater detail. (For the purposes of the discussion it will be necessary to collapse rationality and prudence, pace Derek Parfit, since legal discussion is almost always in terms of reasonableness or rationality, rather than prudence.) Recall that in Brandt, the most basic criterion for rational decision-making was that all relevant information should be in view. This standard would rule out the possibility of a rational decision for Mrs. Sidaway if, as she claimed, the doctor had not provided her with two

crucial pieces of information: the elective nature of the operation and the .01 probability of damage to the spinal cord as a result. It is the second which raises the question about balancing probabilities and utilities in rational decision strategies. Given that Mrs. Sidaway might reasonably have attached a higher "disutility" to paralysis ($p=.01$) than to nerve end damage ($p=.02$), could doctors satisfy their patients' informational requirements for informed consent if they used an expected-value strategy for deciding which risks to disclose? Or does the patient need to know all risks? This apparently impossible demand is only the least stringent of Brandt's requirements for rational decision-making, remember.

It is naive to believe that all risks can be known. Whether or not this is true of Brandt's example of the car which turns out to be a "lemon" even though the purchaser has dedicatedly perused Consumer Reports, it is unarguable, I think, in new therapeutic or experimental medical procedures. This is both a practical and a theoretical caveat. If the doctrine of "informed" consent is construed as requiring all risks to be disclosed, it cannot satisfy the Kantian "ought implies can" standard. The problem is less severe if the doctor's duty is not to impart information about all possible risks, but about all the risks he knows.²⁶ Provided that this does not become the lazy doctor's excuse for failing to keep abreast of research, this reformulation should satisfy a deontologist.

However, the medical doctrine of therapeutic privilege is invoked by some consequentialists to deny that health care personnel should reveal all risks even if they could. Doctors profess a fear of "information overload"

in patients; they sometimes assert that their obligation of primum non nocere (first do no harm) prevails over telling the full truth if the truth could harm, as in the earlier example of a patient who might be driven to suicide by being told she has an incurable disease. This principle was stated for cases of informed consent in Salgo v. Leland Stanford Jr.:²⁷

[There are two possible alternatives:] one is to explain to the patient every risk attendant upon any surgical procedure or operation, no matter how remote; this may well result in alarming a patient who is already unduly apprehensive and who as a result refuses to undertake surgery in which there is in fact minimum risk; it may also result in actually increasing the risks by reason of the physiological results of the apprehension itself. The other is to recognize that each patient presents a separate problem, that the patient's mental and emotional condition is important and in certain cases may be crucial, and that in discussing the element of risk a certain amount of discretion must be employed.

The most extreme form of this argument actually asserts that consent forms constitute an iatrogenic harm to the patient. Knowing all the risks can damage your health, it seems--although some proponents of this view often assert simultaneously that nobody reads consent forms and that nobody will consent if the form lists too many risks. The claim is all the more presumptuous because its best-known authors, James Fries and Elizabeth Loftus,²⁸ intend it also to apply to subjects in non-therapeutic experiments, who by definition are receiving no countervailing medical benefits. Fries and Loftus present a contradictory picture of patients' and experimental subjects' rationality. On the one hand they depict those on the receiving end of treatment as volatile, hysterical, and suggestible; on the other, they suggest that even these benighted souls could understand risks presented in the form of comparisons ("as dangerous as an airplane trip") rather than conventional probabilities. This only goes to undermine

their previous consistent stand against disclosure, since research has shown that people find comparisons, if anything, less informative than absolute levels of risk.²⁹

Nevertheless, there is a sizeable gap between the number of risks which doctors normally reveal to patients and the extent of risk information which patients profess to desire. A 1981 American survey discovered that doctors said they usually disclosed five or six out of a possible sixteen relevant facts about the probabilities of adverse side effects from seizure medication.³⁰ But patients and their relatives wanted to know an average of thirteen or fourteen such facts. Are patients more rational, in Brandt's first sense of requiring complete information, than health care personnel take them to be? It would be pleasant to think so, but other studies suggest that they fail to make fully rational use of information--to measure up to Brandt's second criterion for rationality, the elimination of cognitive bias.³¹ That is not an automatic argument for withholding information from patients, although some writers have fallen prey to the naturalistic fallacy by claiming that because people fail in fact to exhibit perfect understanding of information, there is no obligation in morality to inform them.³²

The dilemma of inefficient use of information is a genuine one for a consequentialist, however. A deontologist might assert that the health care professional's duty is only to inform, not to ensure comprehension. I find this a plausible way of limiting agents' responsibility. Of course, it could become a convenient excuse for peremptory and obfuscating

explanations, but it does at least separate the action of informing from the outcome, skewed by near-ineradicable "availability" and "anchoring" biases.³³ The effect of these cognitive short-circuits is evident in one study of cardiac patients.³⁴

Twenty men with a previous history of hospitalisation for heart disease and with forthcoming operations were informed, in a taped interview before surgery, of the risk of death during the procedure, possible complications, likely benefits, probability of overall success, and alternative treatments. When interviewed again after the operation, only 25 per cent remembered that the possibility of death during surgery had ever been mentioned. A full 90 per cent had forgotten that less serious complications had also been discussed. Even when the interviewers told them that these topics had been included, 58 per cent could not remember the probability of death mentioned, and 77 per cent failed to recall either the content or the probability of other complications. Patients were best able to remember the benefits which had been predicted from a successful operation--indicating contamination by motivational biases, wishful thinking. Each of the patients was unable to recall some important component of the interview, and sixteen out of twenty positively denied that certain major items had come up. Thirteen out of twenty fabricated facts. One patient maintained stoutly that he had been given no information at all: "All he [the doctor] did was lift up my shirt, put a stethoscope on my heart, and that was it." This patient had been given about half an hour of the surgeon's time, the normal length of interview in the study.

As if this were not sufficient indication of irrationality, all the patients thought they knew more than they did: "They were frequently in error but never in doubt."³⁵ The patient who achieved the poorest recall scores was adamant that he could call to mind every word of the interview. Now all these patients had been "educated" by previous experience of hospitalisation; none was an emergency admission; none was a minor or a mental patient. All had convalesced successfully, which leads to the speculation that hindsight bias was at work: they had "known all along" that the operation would be successful--because it had been successful. The converse form of hindsight bias might apply in consent cases which come to litigation. Since by definition these procedures have ended badly, plaintiffs would be likely to overestimate the doubts which they had at the time of signing the consent form. Perhaps, too, they might blame doctors unfairly for not giving them more information to assuage the doubts they probably never had.

The doctors conducting this study concluded that patients were well informed before the operation, but forgot. However, there seems to have been no test of their comprehension at the time. This is crucial, since "informed" consent means informed at the time of signing the consent proforma. In the absence of actual tests, it might be safer to assume that the patients never fully understood the information. If so, was it a waste of the surgeons' time to inform them?

A utilitarian cost-benefit analysis might well conclude from this research that informed consent is hardly worth the cost. However, there are

some opposing findings. It may be no coincidence that all twenty patients in this study recovered: other researchers have found that patients given more information show better rates of success in treatment.³⁶ Another study indirectly explains why this should be so by confirming that patients value a full presentation of statistical information before signing the consent form.³⁷ Perhaps there is some holistic correlation between being valued--between respect for persons, in fact--and responding well to treatment. If this seems too speculative, it is nevertheless fallacious to focus too narrowly on patients' inferior understanding of probabilities, as compared with doctors' comprehension. Practitioners writing for practitioners have asserted that doctors are no better calibrated in assigning probabilities of success than patients are.³⁸

American law has tried to bridge the gap between patients' expressed desire for full information and their unimpressive track record with the information which they are given by developing a doctrine of "material risks." This is the requirement that before obtaining consent, the health professional must "explain the procedure to the patient and...warn him of any material risks of dangers inherent in or collateral to the therapy,"³⁹ not of all risks. But this concept lands us back in the complexities of what constitutes rationality: a material risk is defined as one which "a reasonable person, in what the physician knows or should know to be the patient's position, would be likely to attach significance to."⁴⁰ To the extent that an English law of informed consent exists, it also uses the reasonable person criterion, urging the health care professional to picture a person of average competence with statistics and medical facts in

deciding how much to disclose.⁴¹ This leads to particular difficulty in cases of consent involving the mentally ill, those with learning disabilities, and children. Such instances are worlds away from the typical "rational" individual in Brandt's examples, the Consumer Reports-reading university professor.

A particularly poignant case⁴² involved a woman with arteriosclerosis of the brain who was alternately rational and confused. When confused, she was prone to wander. On relatives' urging, she was hospitalised without her consent by order of a district (lower) court. The woman brought an appeal, with her lawyers maintaining that the burden should be on the lower court to suggest alternative treatments for her wandering. When ordered to reconsider by the court of appeals, the district court held that in her lucid moments, the woman would have seen that the rational decision was hospitalisation. The rational self would have elected to hospitalise the mad self, and itself as a hostage. The woman was returned to the mental hospital, where she died.

This seems to be a genuine and tragic case of being asked to treat one's self at another time as another person, of the sort developed in hypothetical instances by Parfit. In less gothic but equally *Catch-22*-ish cases, senile and otherwise confused patients are judged to be rational if they consent to a procedure. But refusing treatment is proof of irrationality, since treatment is clearly in one's best interests.⁴³ The extreme utilitarian slant of this view is consistent with the obverse cases in which treatment which a mental patient has actually requested is

withheld, often for reasons of cost and benefit. In another case centring on an involuntarily institutionalised mental patient, an American state circuit court refused to allow an experimental procedure: implanting electrodes and perhaps progressing to psychosurgery in order to control aggression and violent sexuality in a male resident.⁴⁴ By the time of this case, psychosurgery was coming under heavy fire, even though its originator had earlier been awarded the Nobel Prize. The court doubted that the benefits of treatment would outweigh the risks, and overruled the patient's consent to the procedure proposed by the doctors.⁴⁵

The "reasonable person" criterion may set unrealistic standards for "normal" patients, too. If the guideline is the patient of ordinary competence with medical facts and statistics, one would only need to recall one risk in three, minus a standard deviation, in order to fall within the normal band suggested by the cardiac patients study. This seems a very low standard of rationality; yet the "reasonable person" is supposed to be the average person. I would suggest that the "reasonable person," despite the concept's claim to be rooted in empirical fact, is in truth a moral concept, not an "is" but a "should." In this sense it is impossible to treat the concept as a merely prudential one, and informed consent decisions as practical matters of ascertaining what risks the "average" "rational" person would want to know if interviewed before signing a consent form. "Average" is a statistical concept, "rational" a hypothetical or ethical one. Consent cases cannot be hived off as prudential rather than ethical.

To summarise, it seems to me that the paradoxes of rationality and risk require an absolutist interpretation of consent. The doctor's responsibility is not to get the outcome right in all cases: it appears doubtful that even a very extreme consequentialist could argue that. Here the normal standards of negligence apply: the doctor will not be found negligent if he performs at the level of average competence for a health professional in his circumstances, which may include extreme overwork or other harbingers of outcome ill-luck. Of course, if the only doctor in town is habitually drunk during half his operations, and tells me so beforehand, I can hardly be said to consent freely if I allow him to operate on me for a fatal condition so as to have at least a 50-50 chance of life. But one hopes he is below the normal clinical standard there, and that he would be liable for negligence even if I had given my consent. In effect, his negligence would have changed the probabilities: they have become very much worse than those I could have expected. Another way of putting this point is to say that although the giving of informed consent does transfer responsibility to the patient, the doctor is still subject to the normal professional standards.

If a procedure turns out badly, the normally competent physician cannot be held ethically at fault if he has obtained informed consent; to put the matter in deliberately over-simplified terms, he will be unlucky but not evil, and he should experience regret but not guilt. The doctor's responsibility is not necessarily to get the outcome right, but to proceed in the correct fashion. I would argue that the second belongs to the realm of ethics, the first to that of prudence, probability and rationality. It

is the consent proforma which separates the two. A paternalist of the extreme sort, by contrast, would seem to be required to regard the doctor as a wrongdoer if the operation turns out badly. Thus an absolutist interpretation of consent protects both doctor and patient: the doctor from moral luck, the patient from invasion of autonomy.

FOOTNOTES

1. Carolyn Faulder, Whose Body Is It? The Troubling Issue of Informed Consent (London: Virago, 1985), p. 2.

2. There is another possibility: the question may not be whether informed consent transfers responsibility to the patient, but whether it diverts it to the law. Many American doctors do not see a signed consent proforma as protecting them--they may still be sued--but merely as bringing in another party with an interest in or responsibility for an unfavourable outcome: the lawyer.

3. I have deliberately over-simplified this contrast. Especially in the United States, the line between consent and negligence cases is becoming very muddled. A consent case is now, typically, not a civil action for assault, but for negligence in giving insufficient information although a formal consent was obtained. See Frank Harron, John Burnside and Tom Beauchamp, Health and Human Values: A Guide to Making Your Own Decisions (New Haven: Yale University Press, 1983). In English law the deciding case is *Chatterton v. Gerson*, 1980, summarised in Faulder, Whose Body?, p. 53.

4. Harris, Violence and Responsibility, *passim*.

5. Harris, Value of Life, pp. 205-212. Tom L. Beauchamp and James F. Childress (in Principles of Biomedical Ethics, second edition, Oxford University Press, 1983) consider that the latter case is difficult even within a single, Kantian framework because principles of beneficence must be balanced against those of truthfulness and respect for the autonomy of others.

6. Ibid., p. 212.

7. Schloendorff v. Society of N.Y. Hospitals, 211 N.Y. 1125, 127, 129.

8. Natanson v. Kline, Supreme Court of Kansas, 1960, quoted in Faulder, Whose Body?, p. 13.

9. Canterbury v. Spence, Federal Republic, 1972, quoted in Faulder, Whose Body?, p. 14.

10. State of Massachusetts statute, 1979, quoted in Faulder, Whose Body?, p. 14.

11. See M.D. Kirby, "Informed Consent: What Does It Mean?", Journal of Medical Ethics, v. 9, no. 2 (June 1983), pp. 70-71: "Usually, of course, claims arise only when something has gone wrong...Indeed, usually unless something seriously wrong has occurred, the costs, delays and other inconveniences of litigation dissuade patients and their families from suing."

12. Again I have deliberately over-simplified, for the sake of contrast. A 1974 article by Jeffrie Murphy (quoted in Beauchamp and Childress, p. 75) remarks that about 75% of U.S. jurisdictions then obeyed a standard much closer to the prevailing English criterion, Bolam: a doctor is only obliged to divulge as much as risk as the average competent doctor would reveal. However, since 1974, the tendency in American law has been very much away

- from this English approach of letting the profession set its own standards.
13. 502 P.2d 1, decided 27 October 1972 in the California Supreme Court.
14. Decided by the Law Lords on 21 February 1985, reported in the Guardian 22 February.
15. In fact, at the Court of Appeal level, Lord Justice Dunn had rejected the concept of informed concept utterly, as being "no part of English law". This view was seconded by Lord Justice Browne-Wilkinson, who warned against extending the concept anywhere "outside the field of property rights" (quoted in Faulder, Whose Body?, p. 55). The Law Lords Scarman and Templeman, in contrast, indicated their support for the development of the concept, but held that Mrs. Sidaway had failed to prove her case.
16. Faulder provides a cogent summary of the anti-paternalist case, at p. 54 and passim.
17. Why we should value full disclosure cannot be answered within consequentialism, however, just as the values of fairness and equal distribution had to be imported into consequentialism in the case of scarce medical resources (see chapter five). The "medical standard" approach follows from the decision in Bolam v. Friern Hospital Management Committee (1957), in which a doctor defeated an allegation of negligence by satisfying the court that his method was one of two standard practices. The court declined to adjudicate between the two, on the grounds that it possessed no medical expertise, and indeed in this decade the House of Lords has ruled that a judge is not allowed to prefer one body of medical opinion to another in his holding. In Sidaway Lord Scarman denied that Bolam was an appropriate model, being a case of negligence rather than consent, but he declined to substitute another.
-

18. Diana Brahams, "To Be Perfectly Frank," Guardian, 25 June 1985.
19. Williams, "The Self and the Future," in Problems of the Self and "Persons, Character and Morality," in Moral Luck.
20. Parfit, Reasons and Persons, p. 320.
21. See a modified exposition of this view in John Kleinig, Ethical Issues in Psychosurgery (London: George Allen and Unwin, 1985), p. 19, and a more extreme version enunciated by Mark S. Komrad in "A Defence of Medical Paternalism: Maximising Patients' Autonomy," Journal of Medical Ethics, v. 9, no. 2 (March 1983), pp. 38-44. A particularly frightening version of the view in favour of overriding refusal to consent to ECT is to be found in Richard Sherlock's contribution to the symposium on consent, competency and ECT in the Journal of Medical Ethics, v. 9, no. 3 (September 1983).
22. Case 4 in Beauchamp and Childress, Principles of Biomedical Ethics.
23. Ibid., pp. 288-289.
24. Case 7, ibid., p. 291.
25. Day seminar in medical ethics for nursing staff, John Radcliffe Hospital, Oxford, March 1987.
26. This qualifier is suggested by Alan Donagan in "Informed Consent in Theory and Experimentation," Journal of Medical Ethics, v. 2, no. 4 (1977), 310-327.
27. 154 Cal. App. 2d 560, 317 P2d. 170.
28. James F. Fries and Elizabeth Loftus, series of letters in exchange with Ruth Barcan Marcus, Bruce Kuklick, and Saevan Bercovitch in Science (1979), reproduced in Gorovitz et.al., Moral Problems in Medicine, pp. 187-191.
29. Research by Sarah Liechtenstein et.al., summarised in block 3 of course U201 (Risk) (The Open University, 1980).

30. Ruth Faden et.al., "Disclosure of Information to Patients in Medical Care," Medical Care, no. 19 (July 1981), pp. 718-733, in Harron, Burnside and Beauchamp, Health and Human Values, p. 91.
31. Brandt, Good and Right, p. 11 ff.
32. See, for example, Franz J. Ingelfinger, "Informed (but Uneducated) Consent", New England Journal of Medicine 287 (31 August 1972), pp. 465-466, in Gorovitz et.al., Moral Problems in Medicine, pp. 179-181.
33. See Kahneman, Slovic and Tversky (eds.), Judgement under Uncertainty.
34. Robinson and Merar, "Informed Consent: Recall by Patients Tested Postoperatively", in Gorovitz et.al., Moral Problems in Medicine, p. 182 ff. A similarly depressing study of consent form recall has been made for cancer patients (Cassileth).
35. Ibid., p. 184.
36. Interview with Barbara Vaughan, Senior Nurse, Nursing Practice Development, Oxfordshire.
37. Ralph Alfidi, "Informed Consent: A Study of Patient Reaction," Journal of the American Medical Association 216 (1971): 1325-1329.
38. Alastair V. Campbell and Roger Higgs (In That Case: Medical Ethics in Everyday Practice, London, Darton, Longman and Todd, in association with the Journal of Medical Ethics, 1982) believe that doctors have a "restricted and inflexible understanding of scientific probability" (p. 28). In particular, doctors feel ignorant of relevant probabilities, are uneasy and overawed about using the ones they do know, and have difficulty assigning a probability in an individual case.
39. *Sard v. Hardy*, 379 A.2d 1014, 1020 (Maryland, 1977).
40. *Canterbury v. Spence*, emphasis added. See also *Wilkinson v. Vesey*, 295

A.2d 676 (1972), Cobbs v. Grant, and Berkey v. Anderson, 82 Cal. Reporter 67.

41. J.K. Mason and R.A. McCall Smith, Law and Medical Ethics (London: Butterworth, 1983), p. 121.

42. Lake v. Cameron, 267 F.Supp. 155 (D.D.C. 1967), summarised in Beauchamp and Childress, Principles of Biomedical Ethics, as case 9.

43. Many commentators agree that this is common practice. See: Loren H. Roth, Alan Meisel and Charles W. Lidz, "Tests of Competency in Consent to Treatment," American Journal of Psychiatry, v. 134, no. 4, 279-284 (1977) in Gorovitz et.al., Moral Problems in Medicine, pp. 172-179; Faulder, Whose Body?, p. 141; and Charles M. Culver and Bernard Gert, Philosophy in Medicine (Oxford: Oxford University Press, 1982).

44. Kaimowitz v. Michigan Department of Mental Health, Civil Action 73-19434-AW (1973).

45. The obvious question is how the case got into court in the first place, if the patient and the doctors were in agreement. The suit was brought by the Michigan Medical Committee for Human Rights as a test case, rather than directly on the patient's behalf. Once the patient was freed--the court also having found that he had been held unconstitutionally for eighteen years--he immediately withdrew his consent.

Chapter Eight

The Illest Chance?

Luck, Moral Mathematics and Nuclear Deterrence

Nuclear deterrence is the decision which we have to "get right" this time. This claim may seem so obvious as to be a mere platitude. If it is, however, it has not been equally evident to strategists, about forty per cent of the adult population¹, or many professional philosophers that deterrence is inadmissible because it stands a chance of "getting it wrong." In order to be credible, nuclear deterrence must maintain that the probability of nuclear holocaust is greater than zero. This holds whether nuclear conflagration is accidental or deliberate. The decision to build up a nuclear arsenal itself creates the possibility of accidental disaster and sets the probability juggernaut rolling. Committed advocates of deterrence might reply that there is absolutely no chance of accident, but it is hard to see how anyone could seriously defend that claim. Even if it were admitted--which I do not mean to do--deterrence rests on a publicly stated willingness to consider deliberate use of nuclear weapons. The decision deliberately to use these arms would of course require policy makers to entertain in their minds a probability of use greater than zero.

Because atomic holocaust is the nearest parallel in a secular age to Pascal's idea of an infinite loss--the eternal destruction of hellfire--the probability of that loss is intolerable if greater than zero, no matter how minuscule. As Barrie Paskins and Michael Dockrill put it:

The precise probability is beside the point...This seems to us an important finding whose significance is understandably underrated

in the strategic tradition.²

This Pascal's wager style of argument against the nuclear deterrent is linked to a reasoning based on outcome luck. Once the likelihood of tragic loss is admitted to be greater than zero,

[r]eliance on the deterrent imports into the lives of those who man the deterrent, and the people in whose name they act, a deep and systematic reliance on luck...[T]rusting to luck becomes the foundation of the security on which civilised life is built.³

Not only is the fate of nations allowed to depend on luck by a policy of nuclear deterrence: so is the policy-maker's own power of decision. It would be a peculiar piece of moral self-indulgence to see ethical agency as more important than the fate of nations; nevertheless, the evaporation of the power of moral choice should be profoundly troubling, and not only to Kantians. As Paskins and Dockrill remark, proponents of nuclear deterrence often take comfort in the assertion that deterrence rests not on the conditional intention to retaliate, but on a bluff. But strictly speaking, bluff also involves a conditional intention--to do the opposite of what I proclaim I will do, to refrain from retaliating if that is what I have threatened to do. What prevailing nuclear strategy actually rests on is a conscious lack of intention, a deliberate cultivation of akrasia, a sort of wilful willessness. If charged with the evil intention of annihilating civilians, proponents of mutual assured destruction often counter that they have no such definite intention: those with their fingers on the button will make up their minds on the spur of the moment.⁴ Why this should be reassuring is hard to fathom.

I shall return to this large and complex dilemma at the end of this chapter. It represents the most profound sense in which denying the existence of the moral luck paradox could undermine the ethical enterprise, I think. Not only the outcomes of decisions would depend on chance, but also character; and not only agents' character, but their power of making ethical choices in the first place. To see this actually presented as an ethical argument for nuclear deterrence is profoundly ironic. If there does not remain even the good will, or any will at all, what does remain of choice and agency?

The study of nuclear weapons policy also fleshes out other points from my first four chapters, particularly those dealing with "moral mathematics", rationality and uncertainty. I shall begin with the last concept, drawing particularly on the work of Robert Goodin in developing a consequentialist defence of unilateralism. (Deontologists have no monopoly on opposition to nuclear deterrence; indeed, there is a sort of consequentialist unilateralism which argues that holocaust would be such an overwhelming loss that we should spend all our days protesting against governmental deterrence policy.⁵) Goodin states that even if we cannot assign particular probabilities to nuclear risks, we are not faced with complete uncertainty. This discussion will then lead into questions of moral mathematics and rationality.

Whether nuclear deterrence is a matter of prudence or of ethics will also have to be considered. This distinction is of course denied by those who believe the rational and the ethical to be synonymous, and a leading

exponent of this school, David Gauthier, argues that deterrence is ethical because rational. Conversely, nuclear deterrence is said by John Finnis and his colleagues to be profoundly irrational and deeply immoral, although, in their deontological framework, it is not immoral only because it encourages irrationality. Both camps consider intentions rather than actual outcomes: consequentialists in this field, such as Gauthier and Gregory Kavka, increasingly focus on the aims rather than the effects of deterrence, even though that might appear to be deontologists' terrain. Holding the intention of nuclear retaliation, however, increases the probability that nuclear war will be the actual outcome, according to Jonathan Glover.⁶ Does a policy of nuclear deterrence thus impose unjustifiable risks, or would unilateral disarmament entail more? Whether it is unethical to impose certain risks on domestic and "enemy" noncombatants will be my last topic of discussion, before I return at the close of this chapter to the theme of moral luck.

First, uncertainty. Is deterrence an area in which ordinary rational calculation is impossibly flawed by lack of information about the odds of catastrophe? Would we want to use ordinary prudential reasoning even if we could? There is something obscene about determining the probability of nuclear holocaust through a policy of deterrence as if it were the likelihood of a car being a lemon. Nevertheless, there have been many stabs at it, despite critics' sneers that no attempt can be any better than guesswork. Daniel Frei, for example, has assigned probability values to four separate categories of risk: technical failure, human error, pre-emptive strike, and disaster through nuclear proliferation.⁷ He points out,

correctly, that no matter how low the probabilities for the four categories separately, they must be added together to give the overall likelihood of disaster. Cumulative risks are what must be considered--over several outcomes and across the years. Although Frei assigns a probability of one catastrophe in 5000 years to the highest of the four single risks, the likelihood heightens to one in 100 years under "normal" international conditions when the four separate probabilities are summed, and to one in 20 under "crisis" circumstances. This figure for times of acute tension should be very worrying for those proponents of deterrence (such as Michael Walzer⁸) who rely on the existence of more or less permanent crisis since 1945 in their arguments in favour of existing policy. It is this abiding state of crisis which is meant to justify overriding the normal just-war prohibition against harming civilians, as counter-city deterrence must contemplate.

But why should proponents or opponents of current strategic policy believe Frei's or any other estimates? Ordinary consumer decisions such as the purchase of a car can be made in a prudential fashion because the correct data are available or at least could be made available. Perhaps this statement needs qualification: no matter how carefully I pore over Which Car? or Consumer Reports, I may still wind up with a lemon. Statistical independence sees to that, even if the information in the reports is as accurate as statistical information can be--no matter how carefully calibrated I try to make my choice. But even if perfect forecasting of the particular outcome is impossible in ordinary prudential decisions, uncertainty is palpably more enormous in the case of nuclear

policy. There is no publication called Which Deterrent? The number of models on the market keeps changing without warning, no sooner available than recalled for defects. Experts appear to be unable to forecast the likelihood of major design faults, such as "nuclear winter", or even to agree whether such a thing as that particular design fault could conceivably exist.

It might be asserted that nevertheless we ought to try to be as well calibrated as possible, and that nuclear strategy should be the paradigm for our efforts to put all decision-making on a rational basis. But how can we generalise from our mercifully defective sample of holocausts? Often scholars who have attempted to assign probabilities for nuclear disaster have had to rely on armchair speculation about estimates from various "experts", a sort of second-hand Delphi method.⁹ By this sort of method, Douglas Lackey has calculated the probability of nuclear war over a 30,000-year timespan, producing complicated expected value figures for numbers of deaths.¹⁰ But in a review of Lackey's computational procedures, Avner Cohen dismisses this quest for spurious numerical precision:

In his ambition to make the case against nuclear deterrence as objective and quantitative, Lackey appears to underestimate one of the most distinctive features of the nuclear predicament--its unique element of uncertainty...This uncertainty is not just a matter of contingent ignorance; the ignorance is essential...The utmost irrationality and uncertainty involved in the eventuality of nuclear war shatters everything, including those modes of thinking that made nuclear weapons possible in the first place.¹¹

This is a deeper qualm than the more obvious ones about calibration and probabilities. The nuclear case is inhospitable to "rational" calculation. A consequentialist argument for or against the bomb has to reckon not only with the lack of hard facts or even of reliable probabilities, but also

with our contradictory intuitions about nuclear weapons. The politics of the nuclear age must be radically different from traditional international relations: how else could we have had a state of permanent crisis since Hiroshima, as Walzer asserts? Longstanding ententes, alliances and enmities are the stuff of old-fashioned international politics, but a permanent cold war which justifies targetting millions of "enemy" civilians is something abysmally new. All traditional wisdom about the management of war would melt away once the nuclear brink was crossed, Cohen claims. Similarly, the nuclear deterrent cannot be accommodated in normal prudential thinking about the conduct of foreign policy. "When nuclear war is at stake, probabilities are not enough: we need moral certainties."¹²

Robert Goodin can be seen as attempting to counter this criticism that uncertainty in the nuclear case is so massive as to render consequentialist reasoning inoperable. The incommensurability of nuclear tragedy can in fact be accommodated within prudential mechanisms, he asserts. Unexpectedly, whilst Cohen opts for multilateralism on the grounds that unilateral disarmament introduces too many uncertainties, Goodin is concerned to provide a consequentialist justification for nuclear pacifism. He applauds unilateralists' campaign "to produce modal changes in the possibility of a large-scale nuclear war rather than just marginal changes in its probability. That limited quest for certainties is, I shall argue, eminently defensible."¹³

Goodin agrees with Cohen that at first glance nuclear deterrence is an unsuitable case for utilitarian treatment. To the arguments about lack of

data on which to base probability assessments and the possibility of infinite loss which cancels out any probability greater than zero, he adds useful remarks on how poorly most individuals deal with probabilities. Now this accurate depiction, drawn from Kahneman, Tversky and Slovic's work on judgement,¹⁴ might be countered in terms of the naturalistic fallacy. Just because leaders and voters are inept at making rational judgements about nuclear risks does not necessarily mean that they should not try to become better at it--quite the reverse, perhaps. The enterprise of trying to make sound prudential decisions about nuclear deterrence might even draw its impetus from the need to improve sloppy thinking. But Goodin does not believe we can do very much better: the odds are inherently unknowable.

In short, my complaint against nuclear deterrence is that it amounts to playing the odds without knowing the odds. That constitutes recklessness par excellence. It would be the height of irresponsibility for anyone to wager the family home on rolls of such radically unpredictable dice. Where millions of lives are at stake, that judgement must surely apply even more harshly.¹⁵

After this ringing denunciation, Goodin might be expected to assert the impossibility of justifying or condemning deterrence policy in terms of risk-balancing. But although he condemns the deterrent, he does so from within a consequentialist framework. Goodin's proposal that we adopt a "possibilistic" rather than a "probabilistic" method of calculation constitutes one more in a series of recent attempts to preserve the consequentialist structure whilst conceding to critics the non-quantifiability of much that classical utilitarianism proclaimed eminently quantifiable.¹⁶ The aim is to shore up consequentialism against critics like S.I. Benn, who, Goodin says, "leaps from the observation that probabilities are unreliable where nuclear strategies are concerned to the

conclusion that consequentialistic reasoning is precluded."¹⁷ Specifically, Goodin wants to retain expected-value models of decision-making under uncertainty, even though these normally require probability assessment. For reasons which I find rather unclear, Goodin seems to eschew decision-theoretic strategies which are independent of probability judgements, such as maximin. Retaining a probability weighting but substituting a three-part division into impossible, certain, and possible is enough to enable expected-value calculation to carry on with its business more or less as usual, Goodin feels. The revamped rule of conduct becomes that "a moral agent should make morally desirable outcomes possible or certain (and morally undesirable ones impossible or uncertain) wherever he can;...a moral agent should open up as many possible paths to good outcomes (and close off as many possible paths to bad ones) as he can."¹⁸ The effect of this supposedly new rule is of course very much the same as that of maximin.¹⁹

Thus the moral agent acts by one set of conventional dictates of rationality--although Goodin has spent considerable time in demonstrating just how irrational people's preferences sometimes seem to be. In the end, however, rationality is assumed, and the possibility of ideologues or "evil emperors" dismissed. Goodin's case for unilateral disarmament rests on what he presents as one certainty in an uncertain world:

There is simply no credible scenario by which a nuclear-armed superpower--provided it is at once minimally rational and governed by the standard goals guiding world politics--would, either by accident or by design, be led to launch a full-scale nuclear assault on an opponent armed only with conventional weapons of a merely defensive sort. In a war of conquest, no aggressor strives to destroy its spoils.²⁰

On the balance of probabilities, this may be so, and it is a point which the opponents of unilateral disarmament usually ignore. But Goodin says he is not dealing with probabilities and balances, although that terminology often slips back into his diction.²¹ To him it is an absolute certainty that no superpower could be so irrational as to prefer annihilating its enemy to enjoying the spoils of a nuclear-free victory. I doubt whether there is any ground for such over-confidence even as far as deliberate destruction goes: what of the fundamentalists, alleged to have numbered Ronald Reagan among their cohort, who look forward to nuclear Armageddon as the occasion on which the just will be "raptured" heavenward?

But even if official policy decisions were rational, a rational superpower can still have accidents. Goodin states complacently that the probability of accidents is zero; yet during an 18-month period following installation of new NORAD computers there were 147 false alerts. Four days after a new warning station in Thule began transmissions, warning of a full-scale attack was received--and revealed just in time to have been a radar mirage caused by reflections off the moon. In the autumn of 1979 new American computers identified another phantom attack, supposedly by Soviet missiles which can reach their target in less than ten minutes. US systems prepared to retaliate; four minutes before the predicted impact, analysts discovered that a training tape had been inserted in the computer system by mistake.²²

The holocaust could occur by chance, and it is partly by chance that it has not yet occurred. Once again consequentialism, even Goodin's revised

version, is mistaken in thinking that it can ignore moral luck. Goodin presents one-sided disarmament as a "sure thing", Cohen more conventionally as a gamble. But conversely, Goodin does remind us that deterrence, although the accepted strategy, is very unpredictable in its payoffs. Both Cohen and Goodin seem to view their opponents' strategies as vulnerable to chance but to present their own prescriptions--multilateralism in the first case, unilateralism in the second--as fail-safe. Overall, I think Goodin holds the edge, insofar as a policy of deterrence actively requires the probability of disaster to be greater than zero, whereas unilateral disarmament does not. However, neither can unilateral disarmament guarantee a probability of zero.

If Goodin's unilateralism is the right moral course, its correctness cannot depend on the accuracy of his prediction that nuclear holocaust cannot possibly ensue. We have already seen serious reasons to doubt the accuracy of that prediction, particularly so far as accidents are concerned, but also in terms of deliberate policy when decision-makers cannot be assumed to be fully rational--ideologues, trigger-happy travellers to Armageddon, or non-elected officials whose policies are not open to public scrutiny. The important point here is that we have no way of knowing whether or not those who actually make the decisions are rational. Once again, our uncertainty is greater than we like to think. It has been alleged that the famous "football" which accompanies the American president everywhere and which supposedly contains the codes for unleashing missiles is an elaborate hoax, designed to reassure the public that the decision is centralised in the hands of one accountable official. Military imperatives

require that the decision actually be decentralised to commanders and sub-commanders in the field, in case the President dies in a first strike. Even the identity of those who will actually make the decision to retaliate is rarely known to the electorate, let alone their rational or irrational state of mind.²³

Perhaps banishing probabilities requires Goodin to assign too many events to the "impossible" and "certain" categories of his scheme. Whatever the reason, the particular "technical fix" which he offers seems not to work. The action of chance cannot be weakened by rejigging consequentialism from a probabilistic to a "possibilistic" mode of operation. Another sort of technical fix in a consequentialist argument focuses on the consequences of the intention to retaliate rather than on those flowing from the act of retaliation. Specifically, a policy of deterrence is held to be both rational and ethical by David Gauthier because intending to retaliate is also rational and ethical, even though actual retaliation would be neither. Gauthier's exercise in "moral mathematics" and the opposing arguments by John Finnis and his colleagues constitute the second major area of discussion in this chapter.

In a deep sense, the policy of deterrence appears to ignore the most basic dictates of rationality. It requires us to hold two contradictory beliefs simultaneously: that the bomb could be used and that it could never be used--precisely because it could be used. The secrecy surrounding nuclear weaponry²⁴ breaches the first, minimal condition for rational decision-making in Brandt, possession of as much relevant information as

possible. Like Sidgwick's "Government House utilitarianism," deterrence appears to be a policy which only works provided that the majority is kept in the dark about it. On the practical "national security" level, this is obvious. The number of Cruise missiles installed at Molesworth and Greenham Common has only been made known to the British public, the supposed beneficiary of protection from the missiles, now that the Soviet "enemy" has had to be told, too, through the 1987 INF summit.

It has also been argued that the policy of deterrence encourages the active cultivation of irrationality in decision-makers themselves, particularly over time. If leaders are rational, it could be disastrous for them if their enemies get to hear of it. Applying an argument originally made by D.H. Hodgson,²⁵ Jonathan Glover points out that sanity can be bad for the nation's collective health. The best outcome in a cold war confrontation is for both sides to avoid nuclear holocaust, but Hodgson viewed this as unattainable because the sane leader will prefer as the second-best outcome to limit nuclear devastation to one side only, even if it is the enemy side, and will not retaliate.

...[I]n certain circumstances the sane man's rank ordering will have disastrous consequences if the other side gets to know of it. [This] does not show that to possess the rank ordering is self-defeating, but it does show that, in a balance-of-terror situation, it may be necessary to conceal the rank ordering by making insincere threats of retaliation...It is hard to see how a government could make insincere threats of retaliation without equally deceiving its own population...A further problem is that the deception and propaganda can work too well...[S]ome people, perhaps including the politicians of the next generation, will be convinced...[T]he insincerity may disappear and a genuine threat emerge.²⁶

Now this prescription in favour of honesty as the best policy--or at least against subterfuge as encouraging irrationality in unpredictable ways--could be accommodated within consequentialism, if not in Glover's variant. As I pointed out in chapter six, it is only Kantianism which asserts that honesty is better than any policy. David Gauthier can be seen as arguing that a consequentialist defence of the deterrent can accommodate Glover's points about unpredictability and irrationality. Implicitly, Glover is saying that the expressed intention to retaliate, even if insincere, changes the probability of actual retaliation, making it unacceptably high. Explicitly, Gauthier incorporates the link between intention and action, the manner in which a policy designed with certain probabilities ostensibly in mind shifts the boundaries of those outcomes' likelihood. To Gauthier, the unpredictability which this introduces into utilitarian calculation is not insuperable; nor does he accept that the policy of deterrence is ambivalent or malign towards rationality.

Game strategists such as Thomas Schelling²⁷ have often claimed that it may indeed be meet to put an antic disposition on--that a prudent player might choose to negotiate from the apparent weakness of irrationality. Although mutual devastation is of course a disaster to any sane person, the calculating statesman may derive considerable advantage from appearing lunatic enough not to realise that. Gauthier implicitly draws on Schelling's ideas, with their Orwellian parallel of "strength through weakness," and on Machiavelli's counsel to appear but not be irrational, from time to time. Although massive retaliation would be massively senseless, the intention to retaliate may be highly politic:

It may be utility maximising to form the nonsubmissive, retaliatory

intention; therefore it may be rational to form such an intention; if it is rational to form the intention it is rational to act on the intention; therefore a rational person can sincerely express the intention; therefore another rational and informed person can be deterred by the expression of the intention.²⁸

There are some large claims here, and some broad jumps between them-- particularly from "if it is rational to form the intention" to "it is rational to act on the intention." Schelling, for example, leaves acting out of it. How does Gauthier justify his assertions?

In the traditional deontological condemnation of the intent to do evil, moral opprobrium flows from the contemplated act to the thought of the act. It is wrong to intend what it would be wrong to do. Gauthier reverses the direction:

It may be tempting to suppose that it is rational to form an intention if and only if it would be utility maximising to execute the intention. Instead we argue that it is rational to execute an intention if and only if it is utility maximising to form it.²⁹

Thus, although intention is generally more prominent in deontological argument, Gauthier tries to incorporate it into consequentialism as a tool of prudent policy. The possession of a good will is the standard Kantian defence against outcome ill-luck. My intentions, at least, cannot be fortune's slaves. But to Gauthier, no longer is a good will the only securely good thing in the world. The will to perform an act is to be measured by the same ethical calculus as the act's outcome--in Gauthier's formulation, by the yardstick of utility maximisation. And it is to be measured independently: Gauthier drives a wedge between intention and outcome. No matter how evil the actual destruction of civilian populations might be, there is no intrinsic harm--and possibly much politic benefit--in thinking very seriously about it.

Thinking seriously about doomsday does of course require a doomsday machine. When this is set up, the odds on destruction may well slip from out of the makers' control, as Frei alleges, and moral luck creep in. Gauthier does recognise that calculation is a continuing process rather than a single act in the case of deterrence, that, for example, "adoption of the intention affects one's expected utilities by affecting the probability that the condition for implementation will be realised."³⁰ Nevertheless, this is intrinsically different from a consumer decision or a bet, even if the probabilities there may have to be recalibrated and the utilities reassigned in the course of the decision. The odds which Honest Joe is offering will lengthen and shorten as the race approaches, according to the stake money put down by myself and other punters, but my choice of how much to bet on which horse and at what stage does not affect whether or not the race actually takes place. I have not set the machine going, so there is no question of it subsequently slipping out of my control. Deterrence decisions, again, are profoundly unlike ordinary consumer decisions, but Gauthier deals with the decision to "buy" a doomsday machine much as if it were any other life choice. (It is, of course, at least potentially a death choice.) The example which he offers as a metaphor for nuclear deterrence concerns an academic whose wife attempts to "deter" him from leaving desirable Boston for the wastes of Texas.

In the end Gauthier really offers little more--despite his large claims--than a standard expected-value analysis:

The benefits of deterrent success must always be balanced against the costs of deterrent failure, and only the relevant probabilities of being in the undesirable situation, both with

and without a policy of deterrence, together with an interval measure of utility in terms of which we may calculate the minimum required probability for deterrent success, enables us to calculate the balance of benefits and costs...The cost of acting on the deterrent intention enters, with appropriate probability weighting, into determining whether it is rational to form the intention.³¹

He merely assumes that calculation can apply to intention, though it is hard to see how aims could be quantified in the same manner as outcomes. It may be that Gauthier intends us to measure the consequences of intentions, but even this poses problems. Is it by luck or design, through correlation or random co-occurrence, that deterrence policy has coincided with forty years of a sort of peace?

Gauthier does not so much prove that the rational is also the ethical as assert it: "Morality, in my view, follows rationality."³² Whether or not nuclear deterrence is ethical is treated ab initio as an empirical question, although a difficult one. But what makes it a particularly insoluble mathematical conundrum--to argue for the moment in Gauthier's terms--are the infinity of the possible loss and the chance that it will occur this time. The point about statistical independence does not come up in Gauthier, but it does trouble Frei:

Statistical probabilities do not provide any information regarding the occurrence of specific single events. A probability of one unintentional nuclear war per hundred years may seem negligible; it looks much more menacing if one assumes that, according to this probability distribution, the disaster may happen tomorrow and be followed by 99 years of peace.³³

If we take the problem of future generations seriously, the occurrence of the disaster at the end of the 99 years should trouble us just as much as its happening tomorrow. But either way, Frei is quite correct to remind us that it could happen anytime. Lest this seem hypothetical doom-mongering,

the US space shuttle programme was projected to lose one vehicle in several hundred flights stretching up to the year 2000; the Challenger carnage took place on the twenty-fifth flight. Although the American public was profoundly shocked by this "unlikely" event, there was nothing remotely surprising about it, in terms of statistical independence.

Gauthier criticises unilateralists for their alleged fixation on the prospect of catastrophe:

The argument of the irrationality of deterrence looks only to the costs of deterrent failure. Because there are such costs, it rejects the policy.³⁴

But what I have been arguing is not that it is because deterrence entails costs that a prudential argument for it must be rejected; it is because it entails costs of a sort which render prudential calculation impossible. It is impossible to provide either a prudential justification or a prudential rejection of deterrence.³⁵ We may lose the world: it is hard to know what loss could be infinite if this is not. Expected-value strategy for decisions under uncertainty and the Parfit-Taurek debate on whether it is more evil to harm five people than one both deal in finite numbers. What we face in the nuclear holocaust is the possibility of a loss so vast as to wipe out probability calculations, as Pascal argues is true of the infinite reward of paradise and the boundless loss of hell. From a purely arithmetical point of view, Pascal is correct. The only query is whether we assign an infinite negative value to the destruction of the world. If the cost is infinite, a rational and prudent policy-maker will wish to avoid it, no matter how infinitesimally small its likelihood. This will be the appropriate course for prudential reasoning: to know its own limits.

If the forfeit of the world is not viewed as a strictly infinite loss, it still appears to me to be of a different and incomprehensibly greater order of magnitude than any of the utilities which could be maximised by deterrence, including the prevention of Soviet tyranny. This view is shared, to their credit, by three writers who consider that Soviet domination would constitute enslavement, as I do not. It helps to explain their assertion, contra Gauthier, that the policy of nuclear deterrence is deeply wrong. As G.E.M. Anscombe condemned the Second World War from within the just-war tradition, on the grounds that it would produce mass civilian deaths,³⁶ so Finnis, Boyle and Grisez draw on Catholic theology to propound an equally controversial platform: "immediate and, if necessary, unilateral renunciation of the deterrent strategy in so far as it involves threats of city swapping and final retaliation."³⁷

Finnis and his colleagues marshal and criticise a wide range of consequentialist arguments both for and against the nuclear deterrent. The concern appears to be not only whether the deterrent is permissible or evil, but whether it can be either right or wrong for consequentialist reasons. Against a first set of outcome-orientated claims they restate what Charles Fried calls the Christian challenge to consequentialism: those willing to sell their souls in order to gain the world may lose both their souls and the world.³⁸ In the case of nuclear deterrence, the second part of this warning is literally true. Against those who might appear to be their allies--unilateralist consequentialists like Goodin--they warn that you can't get there from here.

[Prudential analysis] works only when (a) goals are well defined;
(b) costs can be compared with some definite unit (e.g. of money);
(c) benefits can also be quantified in a way that renders them

commensurable with one another, and (d) differences among means, other than their efficiency, measurable costs, and measurable benefits, are not counted as significant. Moral reasoning, however, (a) is not concerned with some definite goal, but with human well-being or the virtuous life as a whole; (b) concerns values and disvalues [sic] which are diverse in kind, and usually non-quantifiable; (c) seeks goods which cannot be quantified in a way that renders them commensurable; and (d) cannot ignore differences in the quality of the means.³⁹

This listing of the differences between the prudent and the right is more sophisticated than Gauthier's blanket assertion that the second follows directly from the first, but it is open to challenge nonetheless. Proposition (a), for example, would be accepted by an Aristotelian, but there is no reason why an act utilitarian should take it on board. The same is true of proposition (d), which is of course a Kantian claim.

Finnis, Boyle and Grisez spurn the various technical fixes suggested by consequentialists in both the pro- and anti-deterrence camps, always on the grounds that the inherent contradictions in moral mathematics are too vast for tinkering. Against Goodin they reiterate that the rightness or wrongness of deterrence cannot be settled by rejigging the boundaries of the probability distribution. Against Gauthier they assert that the direction of the wrongful-intentions principle--that it is wrong to intend what it would be wrong to do--cannot simply be reversed. Now of course there is no intrinsic reason why a consequentialist should worry about the wrongful-intentions principle, just as he or she could choose to ignore propositions (a) and (d) above. But having decided to incorporate intention into his consequentialist argument, Gauthier is open to this objection.

But are Finnis, Boyle and Grisez right to reject moral mathematics altogether? Is this as blanket a claim as Gauthier's simplistic assertion that the moral follows the rational? I have tended to argue along similar lines, at least so far as Gauthier's calculations go--presenting his argument as failing for reasons which a consequentialist cannot ignore, mistakes in moral mathematics concerning statistical independence and the nullifying effect on probabilities which is exercised by infinitely great losses or gains. I have tended also to assert that because the nuclear case involves the infinite, moral mathematics is invalid here. But before proceeding to my third topic in this chapter, what risks if any can be imposed on others in nuclear policy, I shall let the advocates of moral mathematical reasoning have one last attempt to prove their case.

In an article, "Some Paradoxes of Deterrence,"⁴⁰ Gregory Kavka attempts what I am tempted to regard as merely another technical fix, but which he presents as a new decision-making mode. This method, which Kavka calls his "disaster-avoidance principle," yields a pro-deterrent result, as did Gauthier's. It also resembles Gauthier's later procedure in asserting that the morality of an intention depends on the consequences of the intention, not of the act which follows from it.

Considering deterrence as a policy of intentions, then, rather than as a set of actions, Kavka arrives at an interesting predicament. Deterrence is likely to produce a more serious loss--nuclear holocaust--but unilateralism entails a higher probability of the loss which it may result in, Soviet domination. Although Kavka seems to toy with the realisation

that the cost of deterrence could be an infinite loss, which Soviet control is not, he omits Pascal's wager considerations, which would clearly argue against the deterrent. Maximin, which would entail unilateralism, would seem the appropriate strategy, he suggests, for reasons which are less than clear. Expected-value strategy is made to deal with exactly such cases as Kavka's, in which a higher probability and lower utility for one alternative are to be balanced against the opposite in another, and there is no genuine paradox in this case.⁴¹ However, in an argument whose moral mathematical, empirical basis seems at best very doubtful, Kavka goes on to claim that maximin is not appropriate--because we can be quite confident that over the next thirty years holocaust is extremely unlikely.

There are grounds in statistical independence for distrusting Kavka's vanishing-trick. Unless we can be certain that holocaust is impossible over the next thirty years, it could happen tomorrow, as Frei correctly warns us. But Kavka never says that holocaust is impossible, itself an impossible position to defend. Nevertheless he regards it as being of a different level of probability from the worst outcomes entailed by the options in "ordinary decision theory," where, he says, the lowest-ranked consequence is usually quite likely. This is merely asserted, as is the extreme unlikelihood of holocaust, and illegitimately, I think, in order to get the result which Kavka wants: that maximin is not after all the appropriate strategy. Apart from the dubiousness of Kavka's confidence about the unlikelihood of holocaust, he is using factors extrinsic to maximin--probabilities, which maximin deliberately ignores--in order to reject maximin.

What does turn out in the end to be a revised version of expected-value procedure is formulated by Kavka to apply specifically to situations such as nuclear risk--"cases of two-dimensional uncertainty," in which one option entails the greater "disutility" and the other the higher probability of loss. In this case, says Kavka, "it is rational to select the alternative that minimises the probability of disaster occurrence."⁴² This dictates accepting the risks of nuclear deterrence, with a lower probability although a greater loss. Kavka's calculation is of course open to the arguments about infinite loss which I have already rehearsed, but I do not intend to repeat them. What I want to emphasise as I end this section on moral mathematics and deterrence is this: the possibility that the loss which might ensue from deterrence is infinite is matched by our complete ignorance about the probabilities attached to a policy of unilateral disarmament, a much more radical uncertainty than any that Kavka contemplates.

If we have a limited sample from which to generalise in dealing with use of nuclear weapons--Hiroshima and Nagasaki--we have a sample of zero in considering the likely consequences of unilaterally renouncing nuclear weapons. It has never been done, although the Soviets have come closest in the recent INF negotiations. Not even an extreme Bayesian is willing to compute probabilities from a sample of zero, although he or she might be content with a sample of one. Neither Kavka or anyone else can possibly assign a meaningful probability to the occurrence of Soviet domination through unilateral disarmament by a western power--or of accidental nuclear war during such laying down of nuclear arms.⁴³ A disgruntled hard-liner in

charge of a US missile silo about to be rendered impotent might well push the button rather than let the damned Russkis take over. But there is no branch of moral mathematics which could calibrate the probability of that with even the most minimal degree of confidence.

A freak disaster through disarmament could not be interpreted as meaning that unilateralism was wrong from the start so long as unilateralism was not justified only in terms of the consequences. That is the argument from moral luck as applied to nuclear deterrence and disarmament. The argument from uncertainty and ignorance is linked: the consequences are uniquely unknowable in the nuclear case, and a consequentialist argument either for or against utilitarianism cannot be made without conjuring probabilities out of a hat. For practical reasons to do with the limitations of moral mathematics, neither unilateralism nor deterrence can be justified in terms of utilities and probabilities, on a prudential or consequential basis. For reasons to do with moral luck and with the nature of the moral enterprise--to be considered in greater detail at the end of the chapter--we should actually be glad of that.

The example of the silo commander's last stand would doubtless be seized on by proponents of deterrence strategy as an example of the terrible risks to which unilateral disarmament could expose both us and the "enemy." It is now important to consider some matters to do with imposing risks on others. The deontological tradition on war does appear willing to risk terrible mass suffering for the sake of high-mindedness, opening itself up to the charge of moral self-indulgence. Here is an example from

Anscombe's pamphlet opposing the war against Hitler:

It is said that war admittedly produces a number of evil effects, including attacks on civilians, but that these must be balanced against the probable good effects of the war, and if they are outweighed by good, then they can be discounted. It is indeed true that such a balance must be made, but we cannot propose to sin, because that evil will be outweighed by the good effects of the war. That would be to commit sin, however small, for the sake of any good, however great, and if the choice lies between our total destruction and the commission of sin, then we must choose to be destroyed.⁴⁴

Similarly, Gandhi is said to have advised the Jews to commit suicide rather than resist Hitler. If this is the sort of risk which absolutist injunctions impose, how can they be anything but profoundly unattractive? And if unilateralism is motivated by such a deontological ethic, how can it claim to be any less callous towards civilian populations than deterrence strategy?

There is an empirical sort of answer to this criticism: the actual choice may not in fact lie between total destruction and the commission of sin. (Indeed, Anscombe does not assert that it necessarily does: merely that if it did, we would have to choose not to sin.) Michael Walzer has drawn up an interesting argument regarding passive resistance, purporting to show that if anything, it is armed struggle which is more likely to produce both heavy destruction and the doing of wrong.⁴⁵ The probabilities available to us about civil disobedience against invaders are nearly as inadequate as those we possess--or do not possess--about the consequences of unilateral disarmament, for the same reason: no one has tried it. Non-violent resistance has always been seen as the last chance against an invader, to be invoked only if military defence fails. Admittedly there are no examples of an occupying power being ejected by civil disobedience--

unless we count India, which, curiously enough, Walzer does not mention-- but this results from the self-fulfilling prophecy that only military resistance will be effective.

No non-violent struggle has ever been undertaken by a people trained in advance in its methods and prepared (as soldiers are in case of war) to accept its costs. So it might be true [that passive resistance is as effective as military action, and much less costly in lives]; and if it is, we should have to regard aggression very differently from the way we do at present.⁴⁶

The empirical rejoinder to the accusation that unilateralism would impose massive suffering is that we cannot know this for certain. If anything, it is deterrence strategy that willingly risks the ordeal of the innocent, since it requires the probability of holocaust to be greater than zero. And of course it also demands the doomsday machine to make deterrence credible: it sets rolling the juggernaut of risks.

This leads into a second kind of reply to the charge about imposing risks, one based not on the empirical lack of probabilities on which to base a calculation, but on the moral locus of the rightness or wrongness of deterrence policy. Traditional deontological reasoning, of the sort still embodied in the work of Finnis, Boyle and Grisez, held that it was wrong to intend what it would be wrong to do. Since it would be wrong deliberately to kill innocent civilians, it would be evil to intend to do so, as counter-city deterrence policy must. Now consequentialists have tried to open chinks in the armour of this wrongful-intentions principle by focusing not on the evil of actual retaliation but on the extent to which the intention to retaliate if attacked maximises utility and constitutes a good policy. We have seen an example of this in Gauthier.

An intriguing deontological rejoinder to this bifurcation of intention and action is made by Jefferson McMahon, who condemns deterrence strategy not on the usual wrongful-intention grounds but on the basis of risk imposition.⁴⁷ To McMahon deterrence is wrong not because it embraces a wrongful intention, but because it accepts the risk of deliberately using counter-civilian weapons. Thus the deontological premise changes from "It is wrong to intend (even conditionally) that which it would be wrong to do" to "It is wrong to risk doing that which it would be wrong to do, and wrong to support a policy which carries a risk of wrongdoing."⁴⁸ This makes considerable sense in terms of statistical independence, which means that anyone who accepts a risk of doing that which it would be wrong to do must also accept that that which it would be wrong to do could occur this time.

McMahon's reformulation of the traditional deontological position turns the charge of imposing risks against the proponents of deterrence. Not only do they risk holocaust--which unilateralists can also be seen to do, in the case of the silo commander--but they also risk doing wrong. In a democracy, they also risk subverting democratic accountability itself, since doomsday machines are almost entirely exempt from popular control.⁴⁹ McMahon therefore supports unilateralism on the grounds "that any policy of nuclear deterrence which it would be possible for citizens in a democracy to support would involve a risk that nuclear weapons would be used without their authorisation."⁵⁰

This seems a particularly neat move because it incorporates an inherently consequentialist concept--risk and risk imposition. As I

mentioned in discussing Finnis, Boyle and Grisez, there is no reason why a consequentialist should accept that intent must be part of ethics, and in fact many reasons--given the trickiness of the concept and the nature of the utilitarian enterprise--why he or she should not. But McMahon argues against a consequentialist espousal of deterrence on grounds native to consequentialism.

This concludes my examination of uncertainty, moral mathematics and risk imposition in the example of nuclear deterrence. I have tried to make some relevant points about moral luck as I went along, but now I need to draw them together into a more comprehensive conclusion. Nuclear deterrence policy embodies the willing acceptance of moral luck in a particularly telling way, I think. To illustrate this, let me return to Gauthier's argument in favour of the deterrent.

It is conventionally recognised that deterrence actively requires uncertainty--in the opponent's mind, about whether I will indeed retaliate if he attacks me. But it also demands that I be uncertain in my own mind. I cannot be certain that I will not return force with force, for if I am, I may betray my pacifism in my public pronouncements and effectively invite the enemy to attack me in the secure knowledge that I will turn the other cheek. I cannot be certain that I will definitely retaliate for equally politic reasons. My countrymen prefer to delude themselves with the belief that I am "only bluffing," and too much hawkishness will put the wind up them.

Gauthier correctly points out that I cannot know for sure how I will react if attacked. Will my calculation of the appropriateness of retaliation be less than rational?--moved by anger or panic, for example? The link between retaliation and intention in deterrence is thus not as fixed as some deontologists claim, he argues. This seems a plausible argument, although I think McMahon can bypass it by substituting the risk of doing wrong for the intention to do wrong. But leaving that aside and concentrating on intention for the moment, uncertainty does introduce an interesting quirk.

Gauthier therefore prefers to restate my conditional intention to retaliate as my situation of finding myself choosing to retaliate. This is of course quite a good factual representation of what may actually happen in politics, particularly if a segment of the leadership is killed in the first strike and the doomsday machine moves into semi-automatic operation. But as a depiction of ethical choice it raises enormous problems of moral luck. The agent is so alienated from his or her own decision as effectively to have become a mere observer of it. This relates back to Nagel's dilemma of "excess objectivity", and casts into doubt the whole notion of responsible agency.

With Sartrean "bad faith", I announce through following a policy of deterrence that I have no real choice--that I can neither predict nor control my response to enemy attack. My moral decision--and whether to unleash massive destruction on enemy civilians can hardly be described as anything but the most momentous ethical choice--is allowed to depend

entirely on external circumstances. I adopt what Nagel called a "radically external standpoint" towards my own agency and responsibility. (See chapter four.)

There is an alternative version of Gauthier's argument, also pro-deterrent: US strategic doctrine has been said to depend on the nation being seen to have no choice but to "go nuclear" once conventional war begins in Europe, since nuclear firepower will be in the zone of operations.⁵¹ Although Gauthier pictures the decision as radically underdetermined and American strategy sees it as fully determined, the implications for moral choice are the same. There is none, and the moral agent becomes a cipher in both cases. One pole of the moral luck paradox has been pushed to the extreme: agents are no longer held responsible for how things turn out because those outcomes were beyond their control, and indeed because there is no longer any such thing as agents' control. Yet what could illustrate the urgency of the other side of the moral luck dilemma more powerfully than nuclear destruction? Are we really willing to say that no one was responsible if half of humanity should die?

There are weaker versions of this argument about the effect of a policy of deterrence on the character of those employing it. McMahon asserts, for example, that possessing an evil intention--to retaliate against civilians--can corrupt, much as some consequentialists view habitual secrecy as addictive (see chapter six). But such arguments can be accommodated within consequentialism: I merely need to take an extra variable into account, the risk to my own moral character from the various

courses proposed. What cannot be reconciled with consequentialism, or, I would argue, with any ethical doctrine, is the profound threat to there being such a thing as moral choice. Accepting a policy of deterrence does not turn us into calculating machines: it turns us into spinning jennies or cotton gins, machines without the power of rational calculation. The policy of deterrence is hopelessly paradoxical if it is a choice to deprive ourselves of the power of choice--that, or fearful of freedom.

Nussbaum⁵² argues that what is evil about Agamemnon's sacrifice of Iphigenia is not only the murder itself, but also the bad faith by which he persuades himself that killing his daughter was a good thing. In this deeply repellent manner he ignores the moral ill-luck which forces him to cause suffering through whatever choice he makes--to the Greek fleet, becalmed and plague-ridden, if he abstains, and to his child if he obeys the gods' command. There is an uncomfortable parallel here, I think, with the style of pro-deterrent argument which celebrates mutual assured destruction as having brought us peace for forty years.

FOOTNOTES

1. A Gallup poll taken in August 1986 found that 35 per cent of British men and 45 per cent of women interviewed believed that "nuclear arms do keep

you safe." Roughly 53 per cent of men and 47 per cent of women disagreed. Conversely, women were also more likely to favour unilateral nuclear disarmament than men--49 per cent to 39--and the withdrawal of Cruise missiles: an earlier poll, taken in the winter of 1983 just before Cruise arrived, found a 34-point gender gap on this issue. See Beatrix Campbell, The Iron Ladies: Why Do Women Vote Tory? (London: Virago, 1987), pp. 131-132.

2. Barrie Paskins and Michael Dockrill, The Ethics of War (London: Gerald Duckworth, 1979), p. 239.

3. Ibid., original emphasis.

4. This is the "launch-on-warning" posture, in which the decision-maker--far more likely to be a military man than an elected official--makes up his mind instantaneously, relying on warning sensors and computer logic. Another mechanism, the automatic "hair trigger" launch, leaves even less room for agency and intention; an electronic system sets off the launch signal without human intervention. Which control mechanism--if any--underlies US strategic policy is discussed by Robert C. Aldridge in First Strike (London: Pluto Press, 1983).

5. See, for example, Steven C. Patton, "Individual Responsibility, Nuclear Deterrence, and Excusing Political Inaction," in David Copp (ed.), Nuclear Weapons, Deterrence and Disarmament, Canadian Journal of Philosophy Supplement, v. 12, 1986.

6. Glover, Causing Death, p. 267. See also Parfit, Reasons and Persons, pp. 56-90.

7. Daniel Frei, Risks of Unintentional Nuclear War, United Nations Institute for Disarmament Research (Totowa, New Jersey: Rowman and

Allanheld, 1983).

8. Michael Walzer, Just and Unjust Wars (New York: Basic Books, 1977).

9. The Delphi method was developed at the Institute for the Future, where I worked for a short time in 1970. It relies on the collation of estimates from named experts and the gradual elimination of outliers until some sort of norm is reached. Whether this prejudice in favour of majority opinion is a justified method of lessening uncertainty is very much up for grabs, I think. At the time of the Cuban Missile crisis, very few strategic policy experts would have estimated total Soviet ICBM strength at a puny twenty missiles; yet this turns out to have been their total force (Guardian, 30 January 1989). A maverick who offered an extremely low estimate would have been more accurate than the majority of analysts.

10. Douglas P. Lackey, Moral Principles and Nuclear Weapons (Totowa, New Jersey: Rowman and Allanheld, 1984).

11. Avner Cohen, "Lackey on Nuclear Deterrence: A Public Policy Critique or Applied Ethics Analysis?", Ethics, v. 97, no. 2, 457-472, at p. 466.

12. Ibid., p. 467.

13. Robert E. Goodin, "Nuclear Disarmament as a Moral Certainty," Ethics, v. 95, no. 3 (April 1985), 641-658, at p. 642.

14. Kahneman, Slovic, and Tversky (eds.), Judgement under Uncertainty.

15. Goodin, "Disarmament as Moral Certainty", p. 644.

16. See, for example, Samuel Scheffler's "hybrid conceptions" in The Rejection of Consequentialism (Oxford: Clarendon Press, 1982). In a paper delivered at the 1984 conference of the Aristotelian Society, Michael Slote borrowed from economics the concept of "satisficing" to develop a substitute for the norm in consequentialism, maximising value. This was

claimed to be particularly valuable in judgement under uncertainty. In discussion afterwards, Bernard Williams doubted whether it would in fact be possible to retain act utilitarianism, as Slote wished to do, and simultaneously to substitute "satisficing" for maximising of welfare.

17. Goodin, "Disarmament as Moral Certainty", footnote 25, p. 649, discussing Benn's "Deterrence or Appeasement?", Journal of Applied Philosophy 1 (1983), pp. 5-19.

18. Ibid., p. 649.

19. I owe this observation to Alan Ryan.

20. Goodin, "Disarmament as Moral Certainty", p. 653.

21. As, for example, on p. 653, where he discusses the relative probability of use of various murder weapons. Similarly, on p. 651 he brings in probabilities in the context of first strikes.

22. Report by Senators Gary Hart and Barry Goldwater, summarised in Aldridge, First Strike, chapter 9.

23. Paul Bracken, Nuclear Command and Control (New Haven: Yale University Press, 1983).

24. See my chapter six; Sissela Bok, "Distrust, Secrecy and the Arms Race," Ethics v. 95, no. 3 (April 1985), pp. 712-727; Bok, forthcoming book on disarmament and secrecy (mentioned in correspondence, November 1987).

25. In his Consequences of Utilitarianism (Oxford: Oxford University Press, 1967), chapter four, part one.

26. Glover, Causing Death, pp. 266-267, original emphasis.

27. Thomas C. Schelling, The Strategy of Conflict (Cambridge, Massachusetts: Harvard University Press, 1960).

28. David Gauthier, "Deterrence, Maximization, and Rationality," Ethics v.

- 94, no. 3 (April 1984), 474-495, at p. 479.
29. Ibid., p. 483.
30. Ibid., p. 487.
31. Ibid., pp. 485-486.
32. Ibid., p. 494.
33. Frej, quoted in John Finnis, Joseph M. Boyle Jr. and Germain Grisez, Nuclear Deterrence, Morality and Realism (Oxford: Clarendon Press, 1987), p. 210.
34. Gauthier, "Deterrence, Maximization, and Rationality", p. 486.
35. Just as there were consequentialist advocates of unilateralism, there are deontological justifications of deterrence. One reasoning relies on promise-keeping: retaliation is justified as keeping faith with one's population, even if they are mostly defunct--in which case the problem is defined as the classic one of keeping faith with a dead man. Since defence policies and procedures are notoriously exempt from democratic accountability, on national security grounds, I find it difficult to credit that any actual promise has been made to the populace as a whole, let alone to individuals who voted against the government.
36. G.E.M. Anscombe, "The Justice of the Present War Examined," in Collected Philosophical Papers, v. 3, Ethics, Religion and Politics (Oxford: Basil Blackwell, 1981). It is interesting to note that Anscombe was accused of hindsight bias when she opposed an honorary Oxford degree for Truman after the war. Her critics claimed that she could not have foreseen Hiroshima, Nagasaki, and the saturation bombing of German cities; but in this essay, written in 1939, she predicts mass bombing of civilians as an inevitable consequence of the war.

37. Finnis, Boyle and Grisez, Deterrence, Morality and Realism, p. vi.

38. Charles Fried, Right and Wrong (Cambridge, Massachusetts: Harvard University Press, 1978), p. 2.

39. Finnis, Boyle and Grisez, Deterrence, Morality and Realism, p. 252.

This out-and-out condemnation of prudential reasoning is a little disingenuous on the authors' part, I feel: they do use prudential examples such as investment decisions as parallels for the nuclear case. Similarly, although Finnis and his colleagues appear to reject consequentialism utterly in this passage, earlier they have confused matters by postulating that all sound ethical theories must take the consequences into account (on p. 240).

40. Journal of Philosophy v. 75 (1978), 285-302. See also Kavka's "Deterrence, Utility and Rational Choice," Theory and Decision, v. 12 (1980).

41. Kavka may reject expected-value mechanisms because he regards both probabilities and utilities as incapable of precise formulation in the nuclear case. However, his attitude appears contradictory because he is nevertheless willing to make comparisons between the two options in both probability and utility terms--to assign ordinal rankings if not cardinal values. In the article under discussion he sometimes speaks of "harms of roughly similar quantity," but in a debate with Lackey ("Doubts about Nuclear Disarmament," Philosophy and Public Affairs v. 12, 1983, p. 259, fn. 8) he says that two losses are of the same order of magnitude if the larger is not "a hundred or a thousand times as bad as the lesser." There seems to be serious incoherence in these rankings.

42. Kavka, quoted in Finnis, Boyle and Grisez, Deterrence, Morality and

Realism, p. 193.

43. This point is made by Jefferson McMahon in "Nuclear Deterrence and Future Generations," in S. Lee and A. Cohen (eds.), Nuclear Weapons and the Future of Humanity (Totowa, New Jersey: Rowman and Allanheld, 1984).

44. Anscombe, "Justice of Present War", p. 79. The vehemence of this suicidal conclusion is all the more startling because Anscombe is not totally anti-consequentialist, conceding that "a balance must be made" of good effects and bad outcomes. Elsewhere in the article she examines the probable consequences of the war at some length, concluding--inaccurately for most countries, though probably correctly for Britain up until the early 50's--that it would not usher in the prosperity which was predicted for its successful conclusion.

45. Walzer, Just and Unjust Wars, p. 330 ff.

46. Ibid., p. 330.

47. Jefferson McMahon, "Deterrence and Deontology," Ethics, v. 95, no. 3 (April 1985), 517-536.

48. Ibid., p. 535.

49. Thompson, Political Ethics and Public Office. On p. 37, Thompson goes further, presenting the loss imposed by a policy of deterrence as certain, not simply probable: "The damage to the democratic process is not simply a risk; it is continuous and actual...It is not simply another cost to be counted in the calculus of policy making; it subverts the possibility of a society's conducting that calculus. When deterrence diminishes democracy, citizens cannot adequately weigh the deadly risks of the various forms of deterrence under which they might live, or the various steps that their government might take to reduce the reliance on deterrence."

50. McMahon, "Deterrence and Deontology", p. 535.

51. Paskins and Dockrill, Ethics of War, p. 67: "In some cases NATO can credibly threaten to lose control (e.g. threaten to wage a non-nuclear war...whose likely outcome would be a military and political crisis in which the US had little or no alternative but to use nuclear weapons or in which nuclear war would be all but inevitable)." (Original emphasis.)

52. In The Fragility of Goodness.

Bibliography

Aldridge, Robert C. First Strike. London: Pluto Press, 1983.

Alfidi, Ralph. "Informed Consent: A Study of Patient Reaction". Journal of the American Medical Association 216 (1971): 1325-1329.

Andre, Judith. "Nagel, Williams and Moral Luck". Analysis, v. 43 (1983), 202-207.

Anscombe, G.E.M. "The Justice of the Present War Examined". In Collected Philosophical Papers, v. 3, Ethics, Religion and Politics. Oxford: Basil Blackwell, 1981.

----- "Who Is Wronged?" The Oxford Review, no. 5 (1967), 16-17.

Battin, Margaret P. "Age Rationing and the Just Distribution of Health Care: Is There a Duty to Die?" Ethics, v. 97, no. 2, January 1987.

Beauchamp, Tom L., and Childress, James F. Principles of Biomedical Ethics. Oxford: Oxford University Press, 1983, second edition.

Benn, S.I. "Deterrence or Appeasement?" Journal of Applied Philosophy, v. 1 (1983), 5-19.

Berkey v. Anderson, 82 Cal. Reporter 67.

Blum, Laurence. Friendship, Altruism and Morality. London: Routledge Kegan Paul, 1980.

Bond, E.J. Reason and Value. Cambridge: Cambridge University Press, 1983.

Bok, Sissela. "Distrust, Secrecy and the Arms Race". Ethics, v. 95, no. 3 (April 1985), 712-727.

-----Lying: Moral Choice in Public and Private Life. Hassocks, Sussex: Harvester Press, 1978.

-----Secrets: Concealment and Revelation. Oxford: Oxford University Press, 1986.

Bracken, Paul. Nuclear Command and Control. New Haven: Yale University Press, 1983.

Brahams, Diana. "To Be Perfectly Frank". Guardian, 25 June 1985.

Braithwaite, R.B. The Theory of Games as a Tool for the Moral Philosopher. Cambridge: Cambridge University Press, 1963.

Brandt, Richard. "The Morality and Rationality of Suicide". In Rachels, J. (ed.), Moral Problems. New York, 1975.

-----A Theory of the Good and the Right. Oxford: Clarendon Press, 1979.

Broad, C.D. "The Doctrine of Consequences in Ethics". In Cheney, D.R. (ed.), Broad's Critical Essays in Moral Philosophy. London: Allen and Unwin, 1971.

Butler, Joseph. The Analogy of Religion. Ed. William Gladstone. Oxford: Clarendon Press, 1896.

Calabresi, Guido, and Bobbitt, Philip. Tragic Choices. New York: W.W. Norton, 1978.

Campbell, Alastair V., and Higgs, Roger. In That Case: Medical Ethics in Everyday Practice. London: Darton, Longman and Todd, in association with the Journal of Medical Ethics, 1982.

Campbell, Beatrix. The Iron Ladies: Why Do Women Vote Tory? London: Virago, 1987.

Bibliography

- Card, Claudia. "Women's Voices and Ethical Ideals: Must We Mean What We Say?". Review of Kittay, Eva Feder, and Meyers, Diana T., Women and Moral Theory. Totowa, New Jersey: Rowman and Littlefield, 1987. In Ethics, v. 99, no. 1 (October 1988), 125-135.
- Childress, James F. "Who Shall Live When Not All Can Live?" In Gorovitz, Samuel, et.al. (eds.), Moral Problems in Medicine. Englewood Cliffs, New Jersey: Prentice-Hall, 1983, second edition.
- Cohen, Avner. "Lackey on Nuclear Deterrence: A Public Policy Critique or Applied Ethics Analysis?" Ethics, v. 97, no. 2 (January 1987), 457-472.
- Cohen, L. Jonathan. The Probable and the Provable. Oxford: Clarendon Press, 1977.
- Collard, David. Altruism and Economy: A Study in Non-Selfish Economics. Oxford: Martin Robertson, 1978.
- Cooper, John M. Review of Nussbaum, Fragility of Goodness. In Philosophical Review, v. XCVII, no. 4 (October 1988), 543-563.
- Coote, Anna, and Campbell, Beatrix. Sweet Freedom: The Struggle for Women's Liberation. Oxford: Basil Blackwell, 1987, second edition.
- Culver, Charles M., and Gert, Bernard. Philosophy in Medicine. Oxford: Oxford University Press, 1982.
- Davies, Nick, and Black, Ian. "Security's Legitimate Tasks--and How the Guidelines Are Bent". Guardian, "Watchers" series, 19 April 1984.
- De Finetti, Bruno. The Theory of Probability. Chichester: John Wiley, 1974, v. 1.
- Department of Employment, OPCS. Women and Employment. HMSO, 1984.
- Dickenson, Donna. "Juries, Secrecy and Democracy". Chapter 3 of Block 6 Reading Guide, course D308, Democratic Government and Politics. Milton
-

- Keynes: The Open University, 1988.
- Donagan, Alan. "Informed Consent in Theory and Experimentation". Journal of Medical Ethics, v. 2, no. 4 (1977), 310-327.
- Eels, Ellery. Rational Decision and Causality. Cambridge: Cambridge University Press, 1982.
- Ellsberg, D. "Risk, Ambiguity and the Savage Axioms". Quarterly Journal of Economics, v. 75, 643-669.
- Faden, Ruth, et.al. "Disclosure of Information to Patients in Medical Care". Medical Care, no. 19 (July 1981), 718-733.
- Faulder, Carolyn. Whose Body Is It? The Troubling Issue of Informed Consent. London: Virago, 1985.
- Finnis, John, Boyle, Joseph M. Jr., and Grisez, Germain. Nuclear Deterrence, Morality and Realism. Oxford: Clarendon Press, 1987.
- Frei, Daniel. Risks of Unintentional Nuclear War. United Nations Institute for Disarmament Research. Totowa, New Jersey: Rowman and Allanheld, 1983.
- Freund, Paul A. (ed.) Experimentation on Human Subjects. London, 1972.
- "Introduction". Daedalus, spring 1969.
- Fried, Charles. Right and Wrong. Cambridge, Massachusetts: Harvard University Press, 1978.
- Fries, James F., and Loftus, Elizabeth. Correspondence with Ruth Barcan Marcus, Bruce Kuklick, and Saevan Bercovitch. Science, 1979, reproduced in Gorovitz, Samuel, et.al. (eds.), Moral Problems in Medicine. Englewood Cliffs, New Jersey: Prentice-Hall, 1983, second edition.
- Gauthier, David. "Deterrence, Maximization, and Rationality". Ethics, v. 94, no. 3 (April 1984), 474-495.
- Gilligan, Carol. In a Different Voice: Psychological Theory and Women's

Bibliography

Development. Cambridge, Massachusetts: Harvard University Press, 1982.

Gillon, Ranaan. Philosophical Medical Ethics. Chichester: Wiley, 1985.

Glover, Jonathan. Causing Death and Saving Lives. Harmondsworth: Penguin Books, 1977.

Goodin, Robert E. "Nuclear Disarmament as a Moral Certainty". Ethics, v. 95, no. 3 (April 1985), 641-658.

Gravel, Mike (Senator). Introduction to The Pentagon Papers. Boston: Beacon Press, 1971.

Gruzalski, Bart. "The Defeat of Utilitarian Generalization". Ethics, v. 93, no. 1 (October 1982), 22-38.

Hammond, Peter J. "Utilitarianism, Uncertainty and Information". In Sen, Amartya, and Williams, Bernard (eds.), Utilitarianism and Beyond. Cambridge: Cambridge University Press, 1982.

Hampshire, Stuart (ed.) Public and Private Morality. Cambridge: Cambridge University Press, 1978.

Hare, Richard. "Ethical Theory and Utilitarianism". In Sen, Amartya, and Williams, Bernard (eds.), Utilitarianism and Beyond. Cambridge: Cambridge University Press, 1982.

Harris, John. The Value of Life: An Introduction to Medical Ethics. London: Routledge Kegan Paul, 1985.

-----Violence and Responsibility. London: Routledge Kegan Paul, 1980.

Harron, Frank, Burnside, John, and Beauchamp, Tom. Health and Human Values: A Guide to Making Your Own Decisions. New Haven: Yale University Press, 1983.

Harsanyi, John C. "Morality and the Theory of Rational Behaviour". In Sen, Amartya, and Williams, Bernard, Utilitarianism and Beyond. Cambridge:

- Cambridge University Press, 1982.
- Hassner, Pierre. "Kant". In Strauss, Leo, and Cropsey, Joseph (eds.), A History of Political Philosophy. Chicago: Rand McNally, 1972, second edition.
- Hayek, F.A. Economic Freedom and Representative Government. London: Insitute of Economic Affairs, 1973.
- Hodgson, D.H. The Consequences of Utilitarianism. Oxford: Oxford University Press, 1967.
- Holbrook, John H. "Cigarette Smoking and Cardiovascular Diseases". Revised version of American Heart Association Report of Ad Hoc Committee on Cigarette Smoking and Cardiovascular Diseases for Health Professionals, 1977. In Heller, Tom, et.al. (eds.), Coronary Heart Disease: Reducing the Risk, A Reader. Chichester: Wiley, 1987.
- Howard, Nigel. Paradoxes of Rationality. Cambridge, Massachusetts: M.I.T. Press, 1971.
- Inglefinger, Franz J. "Informed (but Uneducated) Consent". New England Journal of Medicine 287 (31 August 1972), 465-466. In Gorovitz, Samuel, et.al. (eds.), Moral Problems in Medicine. Englewood Cliffs; New Jersey: Prentice-Hall, 1983, second edition.
- Kahneman, D., Slovic, P., and Tversky, A. (eds.). Judgement under Uncertainty. New York: Cambridge University Press, 1982.
- Kaimowitz v. Michigan Department of Mental Health, Civil Action 73-19434-AW (1973).
- Kant, Immanuel. Critique of Practical Reason. Tr. Lewis White Beck. Indianapolis: Bobbs-Merrill, 1956.
- Fundamental Principles of the Metaphysic of Morals (Grundlegung zur

Bibliography

- Metaphysik der Sitten). Tr. Thomas K. Abbott. Indianapolis: Bobbs-Merrill, 1949.
- Lectures on Ethics. Tr. Lewis White Beck. New York: Harper Torchbooks, 1963.
- The Metaphysics of Morals, part two. Tr. Mary J. Gregor as The Doctrine of Virtue. Philadelphia: University of Pennsylvania Press, 1964.
- Kavka, Gregory. "Deterrence, Utility and Rational Choice". Theory and Decision, v. 12 (1980).
- "Doubts about Nuclear Disarmament", debate with Douglas Lackey. Philosophy and Public Affairs, v. 12 (1983).
- "Some Paradoxes of Deterrence". Journal of Philosophy, v. 75 (1978), 285-302.
- Kirby, M.D. "Informed Consent: What Does It Mean?" Journal of Medical Ethics, v. 9, no. 2, June 1983.
- Kleinig, John. Ethical Issues in Psychosurgery. London: Allen and Unwin, 1985.
- Komrad, Mark S. "A Defence of Medical Paternalism: Maximising Patients' Autonomy". Journal of Medical Ethics, v. 9, no. 2 (March 1983), 38-44.
- Kuhse, Helga. The Sanctity-of-Life Doctrine in Medicine: A Critique. Oxford: Clarendon Press, 1987.
- Lackey, Douglas P. Moral Principles and Nuclear Weapons. Totowa, New Jersey: Rowman and Allanheld, 1984.
- Lake v. Cameron, 267 F.Supp. 155 (1967).
- Liechtenstein, Sarah, .et.al. "Judged Frequency of Lethal Events". In Journal of Experimental Psychology: Human Learning and Memory, v. 4 (1978), 556-557.

-
- Macintyre, Alasdair. After Virtue. London: Duckworth, 1981.
- Mackie, J.L. Ethics: Inventing Right and Wrong. New York: Penguin, 1977.
- McGrew, Anthony. "National Security as a Challenge to Liberal Democracy". In Paper 14, course D308, Democratic Government and Politics. Milton Keynes: The Open University, 1987.
- McMahon, Jefferson. "Deterrence and Deontology". Ethics, v. 95, no. 3 (April 1985), 517-536.
- "Nuclear Deterrence and Future Generations". In Lee, S., and Cohen, A. (eds.), Nuclear Weapons and the Future of Humanity. Totowa, New Jersey: Rowman and Allanheld, 1984.
- Mason, J.K., and McCall Smith, R.A. Law and Medical Ethics. London: Butterworth, 1983.
- Moore, George Edward. Ethics. London: Oxford University Press, 1912.
- Nagel, Thomas. "Moral Luck". In Mortal Questions. Cambridge: Cambridge University Press, 1979.
- The View from Nowhere. Oxford: Oxford University Press, 1986.
- Nesson, Charles. "Reasonable Doubt and Permissible Inference: The Value of Complexity". Harvard Law Review, v. 92, 1187-1225.
- Nozick, Robert. Anarchy, State, and Utopia. New York: Basic Books, 1974.
- Nussbaum, Martha Craven. The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy. Cambridge: Cambridge University Press, 1986.
- Parfit, Derek. Correspondence with Charles Fried. Philosophy and Public Affairs, v. 9, no. 4 (summer 1979), 393-397.
- "Innumerate Ethics". Philosophy and Public Affairs, v. 7, no. 4 (summer 1978), 285-300.
- Reasons and Persons. Oxford: Oxford University Press, 1984.
-

Bibliography

Paskins, Barrie, and Dockrill, Michael. The Ethics of War. London: Gerald Duckworth, 1979.

Patton, Steven C. "Individual Responsibility, Nuclear Deterrence, and Excusing Political Inaction". In Copp, David (ed.), Nuclear Weapons, Deterrence and Disarmament. Canadian Journal of Philosophy Supplement, v. 12, 1986.

Prior, A.N. "The Consequences of Actions". Proceedings of the Aristotelian Society Supplement, v. XXX, 1956.

Raphael, D.D. Reply to Prior. Proceedings of the Aristotelian Society Supplement, v. XXX (1956), 118-119.

Rawls, John. "The Independence of Moral Theory". Proceedings and Addresses of the American Philosophical Association, v. XLVIII, 1974-1975.

-----A Theory of Justice. Cambridge, Massachusetts: Harvard University Press, 1971.

Richards, David A.J. A Theory of Reasons for Action. Oxford: Clarendon Press, 1971.

Richardson, Jeffrey T., and Ball, Desmond. The Ties that Bind. London: Allen and Unwin, 1978.

Robinson, George, and Merar, Avraham. "Informed Consent: Recall by Patients Tested Postoperatively". In Gorovitz, Samuel, et.al. (eds.), Moral Problems in Medicine. Englewood Cliffs, New Jersey: Prentice-Hall, 1983, second edition.

Roth, Loren H., Meisel, Alan, and Lidz, Charles W. "Tests of Competency in Consent to Treatment". American Journal of Psychiatry, v. 134, no. 4 (1977), 279-284. In Gorovitz, Samuel, et.al. (eds.), Moral Problems in Medicine. Englewood Cliffs, New Jersey: Prentice-Hall, 1983, second

edition.

Russell, Bertrand. Philosophical Essays. London: Allen and Unwin, 1966, revised edition.

Sanders, David, and Dukeminier, Jesse Jr. "Medical Advance and Legal Lag: Hemodialysis and Kidney Transplantation". UCLA Law Review v. 15 (1968), 367.

Sard v. Hardy, 379 A.2d 1014 (1977).

Scheffler, Samuel. The Rejection of Consequentialism. Oxford: Clarendon Press, 1982.

Schelling, Thomas C. The Strategy of Conflict. Cambridge, Massachusetts: Harvard University Press, 1960.

Schloendorff v. Society of New York Hospitals, 211 N.Y. 1125.

Scruton, Roger. Kant. Oxford: Past Masters Series, 1982.

Sherlock, Richard. Contribution to symposium on consent, competency and ECT in Journal of Medical Ethics, v. 9, no. 3, September 1983.

Sidaway v. Board of Governors of Bethlem Royal Hospital, reported in Guardian, 22 February 1985.

Sidgwick, Henry. Methods of Ethics. London, 1874.

Siegler, M. "Confidentiality in Medicine: A Decrepit Concept". New England Journal of Medicine, v. 307 (1982), 1518-21.

Simon, Herbert. Models of Man. New York: Wiley, 1957.

Smart, J.J.C., and Williams, Bernard. Utilitarianism: For and Against. Cambridge: Cambridge University Press, 1973.

Taurek, John. "Should the Numbers Count?" Philosophy and Public Affairs, v. 6, no. 4 (summer 1977), 293-316.

Thompson, Dennis F. Political Ethics and Public Office. Cambridge,

Bibliography

Massachusetts: Harvard University Press, 1987.

Tolstoy, L.N. Anna Karenina. Tr. Rosemary Edmonds. Baltimore, Maryland: Penguin Books, 1954.

Toulmin, Stephen. Foresight and Understanding. London: Hutchinson, 1961.

Trebilcot, Joyce. "Aprudentialism". American Philosophical Quarterly, v. 11, no. 3, July 1974.

U.S. v. Holmes, 26 Fed. Cas. 360.

Von Wright, G.H. The Logical Problem of Induction. Oxford: Basil Blackwell, 1957.

White, Nicholas P. "Rational Self-Sufficiency and Greek Ethics". Ethics, v. 99, no. 1 (October 1988), 136-146.

Walzer, Michael. Just and Unjust Wars. New York: Basic Books, 1977.

-----Spheres of Justice. Oxford: Basil Blackwell, 1985.

Wilkinson v. Vesey, 295 A.2d 676 (1972).

Williams, Bernard. "A Critique of Utilitarianism". In Smart, J.J.C., and Williams, Utilitarianism: For and Against. Cambridge: Cambridge University Press, 1973.

-----Ethics and the Limits of Philosophy. London: Fontana/Collins, 1985.

-----Moral Luck. Cambridge: Cambridge University Press, 1981.

-----Problems of the Self. Cambridge: Cambridge University Press, 1973.

Zimmerman, Michael J. "Luck and Moral Responsibility". Ethics, v. 97, no. 2 (January 1987), 374-386.